PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7:

A61K 38/00, 38/16, C07K 14/00

A1

(11) International Publication Number: WO 00/18421

(43) International Publication Date: 6 April 2000 (06.04.00)

(21) International Application Number: PCT/US99/22608

(22) International Filing Date: 29 September 1999 (29.09.99)

(30) Priority Data: 09/163,047 29 September 1998 (29.09.98)

(71) Applicant: THE UNIVERSITY OF SOUTHERN CALIFORNIA (US/US); University Park, Los Angeles, CA 90007

(72) Inventors: MARKLAND, Francis, S., Jr.; 2804 Elm Avenue, Manhattan Beach, CA 90266 (US). ZHOU, Qing; Apartment A, 224 North Primrose Avenue, Alhambra, CA 91801 (US).

(74) Agent: CHURCHILL, Margaret; Fulbright & Jaworski L.L.P., 29th floor, 865 S. Figueroa Street, Los Angeles, CA 90017-2576 (US). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CONTORTROSTATIN (CN) AND METHODS FOR ITS USE IN PREVENTING METASTASIS AND OTHER CONDITIONS

(57) Abstract

The amino acid sequence of native contortrostatin was used in a cloning strategy to obtain full-length cDNA and deduced amino acid sequences for a contortrostatin precursor. The precursor includes pro-protein, metalloproteinase, and disintegrin (contortrostatin) regions of the multidomain protein. The sequences can be used to produce recombinant DNA molecules which code on expression for contortrostatin proteins, including biologically active variants and fragments. When formulated as a pharmaceutically acceptable composition, the proteins can be used to treat patients by inhibiting disease processes associated with an integrin binding to an integrin receptor.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
ΑT	Austria	FR	Prance	LU	Luxembourg	SN	Senegal
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
ΑZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium .	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	us	United States of Americ
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Vict Nam
CG	. Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	u	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		·
EE	Estonia	LR	Liberia	SG .	Singapore		

1

CONTORTROSTATIN (CN) AND METHODS FOR ITS USE IN PREVENTING METASTASIS AND OTHER CONDITIONS

FIELD OF THE INVENTION

The present invention relates generally to the field of biochemistry and medicine and in particular to the cloning, sequencing, and production of contortrostatin and its precursor.

BACKGROUND OF THE INVENTION

Breast cancer is one of the leading causes of death among non-smoking women 10 and the spread of the disease from the breast to distant sites is a major cause of death in breast cancer patients. At the time of diagnosis over 60% of breast cancer patients will have disease that has spread from the primary site in the breast to some distant site. Spread of cancer to remove sites, e.g. bone, lungs, liver, brain, a process called metastasis, is a characteristic of malignancy and often leads to inoperable disease. 15 Metastasis is the most common factor leading to death from breast cancer. Control of metastasis offers an important avenue for breast cancer treatment. Cancer cells metastasize through the blood or lymph vessels. The first step of metastasis involves the attachment of cancer cells to tissues around the primary site, i.e., to the extracellular matrix (ECM) via cell surface integrins and other adhesion receptors. Integrins mediate cell-cell and cell-substratum interactions and are involved in bi-20 directional signaling that links the ECM with cytoskeletal proteins. Integrins play an important role in the interaction of mammary carcinoma cells with the ECM. In the second step, cancer cells secrete digestive enzymes that degrade the surrounding

15

20

25

WO 00/18421 PCT/US99/22608

2

tissues allowing the tumor cells to invade these tissues. Eventually, the tumor cells enter the blood or lymphatic system where they repeat the adhesion and invasion steps at a distant (metastatic) site. At this remote site, tumor cells induce the formation of new blood vessels (a process called neovascularization), in and around the growing tumor. These new blood vessels supply nutrients to the metastatic tumor and allow it to grow. Treatments that block any of these steps should act to inhibit metastasis.

Integrins on cancer cells play important roles in tumor invasion and spread. They are a family of proteins found on the cell surface of many cell types that mediate interactions between cells, and between cells and their surroundings. CN binds to a specific integrin on the surface of blood platelets, and blocks the ability of platelets to adhere to one another (a process called platelet aggregation). Platelets are small fragments of bone marrow cells that are found in the blood stream. They have both beneficial and harmful activities. Their useful action is to stop bleeding following injury by facilitating the formation of a blood clot. But, under certain conditions they are involved in blocking arteries that supply nourishment to the heart -- an action that can lead to a heart attack. Integrins are heterodimers, composed of α and β subunits involved in cell-cell and cell-substratum interactions. Integrins serve as receptors for extracellular matrix proteins such as fibronectin, fibrinogen, vitronectin, collagen and laminen. Some of these interactions have been shown to be mediated via an Arg-Gly-Asp (RGD) sequence present in the matrix proteins. Both the α and β subunits are required for fibringen binding. For example, one of the members of the superfamily of integrin cell surface receptors is the platelet membrane glycoprotein (GP)IIb/IIIa which interacts with plasma fibrinogen in platelet aggregation. Integrin cell surface receptors have been investigated in the role of platelets in mediating coronary artery thrombosis and rethrombosis in the genesis of acute myocardial infarction [Zucker, M.B., Sci. American 242:86 (1990)]. For platelet aggregation an RGD sequence present in fibrinogen is essential for the interaction with (GP)IIb/IIIa [Ginsberg, M.H. et al., Thrombos. Haemostas. 59:1 (1988)]. Because of its inhibition of platelet aggregation, snake venom has been the subject of various investigations.

A number of proteins purified from venom of snakes of the *Crotalidae* and *Viperidae* families have been found to inhibit glycoprotein (GP)IIb/IIIa mediated

20

25

30

WO 00/18421 PCT/US99/22608

3

platelet aggregation [see, e.g., Huang, T.F. et al., J. Biol. Chem. 262:16157 (1987); Gan, Z.R. et al., J. Biol. Chem. 263:19827 (1988); Yasuda, T. et al., J. Am. Coll. Cardiol. 16:714 (1990); Trikha, M. et al., Fibrinolysis 4 (Suppl. 1):105 (1990); Trikha, M. et al., Blood 76 (Suppl. 1):479a (1990); Holahan, M.A. et al., Pharmacology 42:340 (1991); Shebuski, R.J. et al., Circulation 82:169 (1990); Yasuda, T. et al., Circulation 83:1038 (1991)]. These proteins, classified as disintegrins, are typically disulfide rich. Moreover, all disintegrins isolated thus far, with the exception of barbourin [Scarborough, R.M. et al., J. Biol. Chem. 266:9359 (1991)] contain an RGD (Arg-Gly-Asp) sequence that has been implicated as being involved in the inhibition of integrin-mediated interactions. In particular, the RGD sequence of the disintegrins may compete for fibrinogen binding sites on the platelet membrane, thereby inhibiting platelet aggregation induced by ADP or other agents.

Nonetheless, there appears to be increasing evidence that disintegrins may have unique surface geometry which facilitates interactions with integrins by mechanisms other than those based solely upon the RGD site. For example, the finding that a mutated, chemically synthesized derivative of echistatin (in which alanine was substituted for arginine in the RGD sequence) still possessed some biological activity, suggests that other regions in the protein may be involved in binding and that there may be some flexibility in the RGD binding site [Connolly, T.M. et al., *Circulation* 82 (Suppl. III):660 (1990)]. Synthetic RGD peptides, due to their small size, generally do not possess the molecular topography of the disintegrins and therefore cannot interact via the multiplicity of mechanisms likely to be involved in disintegrin binding.

In other investigations, prevention of reocclusion following thrombolysis using tissue-type plasminogen activator in a canine model system has been reported using 30 μg/kg plus 3 μg/kg/min bitistatin, an 83 amino acid disintegrin derived from the venom of *Bitis arietans* [Shebuski et al., *supra*], or 15 μg/kg/min i.v. echistatin, a 49 amino acid disintegrin derived from the venom of *Echis carinatus* [Holahan et al., *supra*]. In the reported methods, an initial bolus of heparin (100 U/kg i.v.) and subsequent hourly boluses of 50 U/kg were used to increase activated partial thromboplastin times at least 1.5-fold over the control. Whereas it had previously

15

20

WO 00/18421 PCT/US99/22608

4

been observed that heparin in combination with tissue-type plasminogen activator (tPA) did not affect the incidence of acute reocclusion in this model system, the addition of echistatin or bitistatin lead to dramatic reductions in the incidence of acute thrombotic reocclusion. The administration of heparin was, however, apparently necessary for prevention of acute thrombotic reocclusion.

Similarly, kistrin (a 68 amino acid disintegrin derived from the venom of $Agkistrodon \, rhodostoma$) was evaluated in conjunction with recombinant tissue-type plasminogen activator in a canine model of coronary artery thrombosis with superimposed high grade stenosis [Yasuda et al. (1991), supra]. An effective dose of 4 $\mu g/kg/min$ was determined to be sufficient to prevent reocclusion. Simultaneous systemic therapeutic heparin anticoagulation was used; the dose of heparin was selected to maintain the activated partial thromboplastin time more than two-fold throughout the experimental observation period.

U.S. Patent 5,066,592 to Huang et al. describes the use of trigramin, a 72 amino acid disintegrin isolated from the venom of *Trimeresurus gramineus*, to inhibit fibrinogen binding to human platelets and thereby to inhibit fibrinogen-induced aggregation of human platelets. Trigramin is also reported to inhibit binding of von Willebrand factor to platelets. Trigramin is reported to inhibit 125 I-fibrinogen binding to ADP (10 µmolar)-stimulated platelets in a concentration-dependent manner with an IC₅₀ of 2.8-5.6 x 10^{-8} M.

Isolation of an anti-platelet factor applaggin from the venom of Agkistrodon piscivorus piscivorus has also been reported [Chao, B.H. et al., Proc. Natl. Acad. Sci. USA 86:8050 (1989); Savage, B. et al., J. Biol. Chem. 265:11766 (1990)]. Applaggin, unlike trigramin, is reported to inhibit dense-granule secretion in concert with inhibition of platelet aggregation in a dose-dependent manner. While initially described as a homodimer with at least two interchain disulfide bridges [Chao et al. (1989), supra], a subsequent report indicated that analysis of purified applaggin by mass spectroscopy showed the presence of applaggin monomers with a mass of 7,666 Daltons and no evidence of dimerization [Wencel-Drake, J.D. et al., Blood 81:62 (1993)].

15

20

25

WO 00/18421 PCT/US99/22608

5

One disintegrin of particular interest is CN, which has been isolated from the venom of Agkistrodon contortrix contortrix (the southern copperhead snake). The originally-reported purification procedure included molecular sieve chromatography on Sephadex G-100 SF, desalting on Sephadex G-25F and reverse phase HPLC. ADP-enhanced aggregation of stirred human platelet rich plasma and the inhibition thereof by CN were monitored at 37° C. It was found that preincubation for 1 minute of the platelet rich plasma (3 x 10⁵/mm³) with 5 µl of the low molecular weight peak after Sephadex G-100 SF resulted in 76% inhibition of platelet aggregation induced by 10 µM ADP [Trikha et al. (1990), supra].

In a subsequent report it was noted that in crude venom, the inhibitor was not readily detectable due to the presence of platelet aggregating activity; however, following the first step of purification (hydrophobic interaction HPLC) inhibitory activity was separated from both aggregating activity and an α-chain degrading fibrinolytic enzyme present in the venom. Inhibitory activity was pooled following HPLC and applied to a hydroxylapatite HPLC column. In the final step of purification, C4 reverse phase HPLC chromatography was employed. The yield of the homogeneous protein was 3-5 mg per gram of venom. CN was reported to have a molecular weight of 18-21 kDa under non-reducing conditions and 9 kDa under reducing conditions; thus, the molecule was believed to be a homodimer with the two subunits being held together by disulfide bond(s). Isoelectric focusing showed that the protein had an acidic pI. CN was reported not to exhibit fibrinolytic activity and was not a 5'-nucleotidase or a phospholipase based on molecular size and kinetics of inhibition of platelet aggregation. Following preincubation for 1 minute, CN at approximately 100 nM was reported to completely inhibit ADP-induced platelet aggregation [Trikha et al. (1990), supra].

It has further been reported that CN has 70 amino acids with five to six disulfide bridges, and that the sequence of CN appears to begin 10 amino acids downstream of applaggin (a platelet aggregation inhibitor from the venom of *Agkistrodon piscivorus piscivorus*). It was speculated that CN may have an insertion and/or a C-terminal extension of nine amino acids. It was further reported that a 50% inhibition (IC₅₀) of human platelet aggregation in platelet rich plasma was observed

15

20

25

30

WO 00/18421 PCT/US99/22608

6

at 0.8 µg/ml of CN, and at 2.2 µg/ml with canine platelets [Trikha, M. et al., *Journal* of Cellular Biochem. 16F:180 (1992)].

CN was reported to inhibit binding of human fibrosarcoma (HT-1080) and c-Ha-ras transfected rat embryo (4R) cells to fibronectin coated plates but not to matrigel coated plates. Inhibition of 4R cell binding to fibronectin in the presence of CN at 1 μ g/ml and 5 μ g/ml was 46% and 88%, respectively, and for HT1080 cells inhibition was 89% and 85%, respectively [Trikha, M. et al., *Proceedings of the American Association for Cancer Research* 33:34 (1992)].

Since it appears that CN can inhibit interactions between integrins and their receptors, and may prove useful in the management of diseases associated with these interactions, there exists a need for improved methods to produce greater amounts of purified contortrostatin, substantially free other snake venom components.

SUMMARY OF THE INVENTION

The present invention fulfills the need for greater amounts of contortrostatin, which can be used to inhibit biological processes such as platelet aggregation, cell growth, adhesion, metastasis, and neovascularization. Native contortrostatin has been purified and a partial amino acid was determined by Edman degradation. This information enabled a cDNA cloning strategy, which resulted in a full-length cDNA sequence and a deduced amino acid sequence for a contortrostatin pre-cursor protein. The contortrostatin precursor includes a pro-protein region, a metalloproteinase region, and a disintegrin region. The metalloproteinase region includes a metal-binding motif and the disintegrin region includes an RGD loop, which can act as an integrin antagonist. Sequences for native contortrostatin are contained in the distintegrin region.

The present invention includes purified contortrostatin proteins, including the contortrostatin precursor, biologically active variants, and fragments thereof. The contortrostatin protein preferably includes an amino acid sequence that matches native contortrostatin monomer (amino acid numbers 419 to 483 of SEQ ID NO: 2), the metalloproteinase region (amino acid numbers 191 to 410 of SEQ ID NO: 2), the proprotein region (amino acid numbers 1 to 190 of SEQ ID NO: 2) or the contortrostatin

10

15

20

25

7

precursor as a whole (SEQ ID NO: 2). A most preferred purified protein is comprised of contortrostatin monomers, each having a molecular mass of about 5 to about 7 kDa, which form a homodimer.

A purified contortrostatin, which can act as an integrin antagonist generally will include in each monomer a constrained Arg-Gly-Asp (RGD) sequence at the tip of a flexible peptide loop of about 13 amino acid residues flanked by two Cys residues, such as the amino acid sequence comprising amino acid numbers 457 to 469 of SEQ ID NO: 2.

Biologically active variants, can include amino acid substitutions, deletions, and insertions, but will generally have an amino acid sequence that is at least 90% homologous to the pro-protein, metalloproteinase, disintegrin and/or contortrostatin regions of the precursor protein. Variants can also include a peptide recognized by an antibody to contortrostatin.

The proteins of the present invention can be made using synthetic methods. The synthetic process can include transcribing and translating a contortrostatin cDNA molecule as disclosed herein, preferably within a transformed host cell. Alternatively, the process involves synthesizing a polypeptide having the amino acid sequence of contortrostatin, as disclosed herein.

Proteins prepared by recombinant DNA methodology will generally include the use of a recombinant DNA molecule comprising a DNA sequence coding on expression for contortrostatin, such as SEQ ID NO:1. Preferably the recombinant DNA molecule encodes sequences having at least one biological activity such as metalloproteinase (e.g. nucleotide numbers 657 to 1316 of SEQ ID NO: 1), or disintegrin (e.g., nucleotide numbers 1341 to 1535 of SEQ ID NO: 1). Moreover, the recombinant DNA molecule can also include pro-protein encoding sequences (e.g., nucleotide numbers 87 to 656 of SEQ ID NO: 1), the entire precursor protein (nucleotide numbers 87 to 1535 of SEQ ID NO: 1), or simply the sequences encoding native contortrostatin monomer (nucleotide numbers 1341 to 1535 of SEQ ID NO: 1).

The present invention further provides a vector, which includes the recombinant

O DNA molecule, that can be used to transform prokaryotic or eukaryotic host cells. Host
cells can be mammalian cells, plant cells, insect cells, yeast and other fungi or bacteria.

15

20

25

PCT/US99/22608

8 .

Processes for producing recombinant contortrostatin will generally include steps, such as culturing the host cell and recovering the contortrostatin expressed by the host cell.

The contortrostatin proteins of the present invention can be formulated as a pharmaceutically acceptable composition, comprising a pharmaceutically acceptable carrier and the purified protein. The pharmaceutical composition can then be used in a method of treating a patient having a disease associated with a ligand binding to an integrin receptor. The treatment generally involves administering the composition to the patient such that integrin binding to integrin receptors is substantially inhibited and the patient is treated. The method of treatment can be used to inhibit platelet aggregation, tumor metastasis, angiogenesis, neovascularization, cell adhesion, invasiveness, or growth.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention may be better understood with reference to the accompanying drawings, which include:

Fig. 1 illustrates the strategy for cloning contortrostatin cDNA, wherein Fig. 1A shows partial amino acid sequence of CN (CN) based on Edman Degradation Assay compared with other disintegrins illustrating common RGD sequences and highly conserved sequences, Fig. 1B shows the PCR primers, and Fig. 1C shows the principle of overlapping extension reaction used to generate the full-length cDNA of contortrostatin;

Fig. 2 shows electrophoresis of the PCR products, wherein a major band of about 1,300 bp was amplified with PCR primed by λ gt10 forward and PCR-2 (lane 1), a band of approximately 700 bp resulted from PCR primed by PCR-2 and λ gt10 reverse (lane 2), and the overlapping product of the two fragments is shown in lane 3, with a molecular size of about 2,000 bp;

Fig. 3 shows the full-length nucleotide sequence of CN cDNA and the deduced amino acids, including an 86 nucleotide 5'-end non-translatable region (NTR), an open reading frame between nucleotides 87 and 1535, a termination codon at nucleotides 1536 to 1538, and a 3'-NTR, which includes an AATAAA site and ends with a poly-A tail;

Fig. 4 shows the multi-domain structure of contortrostatin precursor, compared with trigramin precursor and other snake venom hemorrhagic proteins;

WO 00/18421

10

15

20

PCT/US99/22608

9

Fig. 5 illustrates the formation of a contortrostatin homodimer, wherein Fig. 5A shows the amino acid sequence and disulfide bond pattern of the disintegrin kistrin compared to contortrostatin, which has a 6-amino-acid truncation at the N-terminus, including two half cystine residues resulting in two cysteine residues being unpaired, and Fig 5B shows the unpaired cysteines may participate in the formation of two intermolecular disulfide bonds to form a unique homodimeric structure, wherein the two monomers may be linked in a parallel or anti-parallel pattern;

Fig. 6 illustrates the results of determinations of CN inhibition of human, canine and rabbit platelet aggregation;

Figs. 7A and 7B illustrate the results of binding studies of CN to human (Fig. 7A) and canine (Fig. 7B) (GP)IIb/IIIa in the presence of a fixed saturating concentration of murine monoclonal antibody 7E3;

Fig. 8 is a schematic representation of an instrumented canine carotid artery showing placement of an ultrasonic flow probe, mechanical constrictor (stenosis) and intra-carotid anodal electrode for inducing intimal injury to the vessel wall to initiate thrombus formation;

Fig. 9 illustrates platelet counts and platelet aggregability as percent of the value at zero time in a canine treated with anisoylated plasminogen streptokinase activator complex (APSAC) and CN;

Fig. 10 shows CN inhibited adhesion of MDA-MB-435 to fibronectin;

Fig. 11 shows CN inhibited adhesion of MDA-MB-435 to vitronectin;

Fig. 12 shows the inhibition of binding of human mammary carcinoma cells to immobilized CN with GRG-DSP;

Fig. 13 shows the inhibition of binding of human mammary carcinoma cells to immobilized CN with EDTA;

Fig. 14 shows the inhibition of invasion of MDA-MB-435 cells through a Matrigel coated invasion chamber;

Fig. 15 shows the effect of CN on the growth MDA-MB-435 tumor in experimental nude mice;

WO 00/18421

PCT/US99/22608

10

Fig. 16 is a photograph demonstrating tumor induced angiogenesis in a control chick embryo chorioallantoic membrane (CAM) (Fig. 16A), CAM treated with 20 μ g of CN (Fig. 16B), and CAM treated with 150 μ g of CN (Fig. 16B); and

Fig. 17 shows the effect of CN on the proliferation of MDA-MB-435 cells in vitro.

5

10

15

20

25

30

DETAILED DESCRIPTION OF THE INVENTION

Characterization, Cloning and Expression of CN

We have purified and characterized a disintegrin: contortrostatin (CN) from $A.\ c.$ contortrix venom. CN is a homodimer with a mass of 13,505 for the intact protein and 6,956 for the reduced and pyridylethylated protein. To test binding affinity to platelet (GP)IIb/IIIa (fibrinogen receptor), competition of CN with [125 1]7E3, an antibody directed to (GP)IIb/IIIA, was analyzed using human platelet rich plasma (PRP). CN displayed an IC₅₀ of 25 nM. Thus, CN is a potent β 3 integrin antagonist.

The cDNA of CN has been amplified from a library of A.c. contortrix venom gland cells, constructed in λ gt10. Amino acid composition and partial amino acid sequences of CN have been determined by Edman degradation; see Fig. 1. Using this information a full-length cDNA of 2,027 nucleotides (SEQ ID NO:1), which encodes a contortrostatin precursor protein, has been cloned and sequenced.

As a member of disintegrin super family, CN shares high similarity with other disintegrins including, trigramin whose nucleotide sequences were known. Fig. 1A shows a partial amino acid sequence of CN based on Edman degradation assay. The partial sequence is also compared with other disintegrins as indicated. The RGD sequence is bold-faced. The highly conserved PCCDAATCKL sequences on which PCR primers are designed are underlined. Figs. 1A and 1B show how the cDNA of CN has been cloned by means of PCR using primers based on the highly homologous sequences among the disintegrin family as well as known \(\lambda\gamma\)10 sequences flanking the cDNA inserts. The PCR primer pairs are: SEQ ID NO:5 (\lambda\gamma\)10 forward primer) and SEQ ID NO:3 (PCR-1) 5'-GATTTACAGGTTGCAGCATCGC-3', which is antisense of trigramin cDNA encoding part of the underlined conserved sequence (Fig. 1A and 1B). SEQ ID NO:4 (PCR-2), which is complementary to PCR-1, and SEQ ID NO:6 (\lambda\gamma\tau\)10 reverse primer). SEQ ID Nos. 5 and 3 amplify DNA coding for amino acids upstream to the underlined part. SEQ

15

20

25

30

WO 00/18421 PCT/US99/22608

11

ID Nos. 4 and 6 amplify those coding for the downstream part of CN. Full length cDNA has been obtained by overlapping extension of the two pieces of PCR products (see Figs. 1C and 2).

The cDNA sequence and deduced amino acid sequence of contortrostatin precursor are shown in Fig. 3. The full length sequence is composed of 2,029 nucleotides (SEQ ID NO:1). It is composed of an 86 nucleotide non-translatable region at the 5'end, an open reading frame coding for 483 amino acids (SEQ ID NO:2), and a 3' non-coding region.

Fig. 4 shows the multidomain structure of contortrostatin compared to that of four other snake venom hemorrhagic proteins: trigramin; Cat (catrocollastatin from *Crotalus atrox* venom); jararhagin (from *Bothrops jararaca* venom); and Ht-e (from *C. Atrox* venom). According to the structural division of snake venom metalloproteinases, the precursor of contortrostatin can be divided into a pro-protein (amino acid residues 1 to 190 of SEQ ID NO: 1 or 2), metalloproteinase (residues 191 to 410 of SEQ ID NO: 1 or 2) and disintegrin (residues 419 to 483 of SEQ ID NO: 1 or 2) domains. The mature monomer of native disintegrin starts at D419, i.e., the aspartic acid residue at position 419. Underlined portions of Fig. 4 show the RGD sequences of both contortrostatin and trigramin, as well as the conserved HEMGHNLGIHH sequences of the zinc-binding motifs in the metalloproteinase domains of each molecule.

According to the present invention, therefore, there is also provided a protein consisting essentially of purified contortrostatin or purified contortrostatin variants retaining the properties of a disintegrin, or a precursor of contortrostatin having proprotein, metalloproteinase, and disintegrin domains. This protein can be purified from natural sources such as snake venom or can be made by recombinant techniques as will be understood by those skilled in the art with reference to the disclosure herein.

Still further, the present application claims both the native and synthetic amino acid and nucleotide sequences. Unless otherwise modified, the term "protein" as used herein encompasses both native and synthetic polypeptides and peptides. Synthetic protein includes recombinant and chemically synthesized protein. Unless otherwise indicated, the term "contortrostatin" include both the native and synthetic versions of the proteins.

WO 00/18421

10

15

20

25

30

PCT/US99/22608

12

The term "nucleotide sequence" includes both the DNA and RNA sequences. For example, the nucleotide sequence for a contortrostatin protein ("contortrostatin nucleotide sequence") includes the gene ("contortrostatin gene") encoding the native and precursor protein, its complementary DNA, and the RNA corresponding to the foregoing; also included are messenger RNA encoding for the contortrostatin protein, its complementary RNA, and the DNA corresponding to the foregoing. Further, as used in this application the nucleotide sequences include: (1) the DNA sequences encoding the contortrostatin proteins, (2) the nucleotide sequences (which may be RNA or DNA) complementary to the foregoing sequences, (3) the corresponding RNA sequences to the DNA sequences wherein the Thymidine ("T") in the disclosed DNA sequences is replaced with Uracil ("U"), (4) nucleotide sequences wherein other nucleotides known in the art such as nucleotide analogs, replace those in the foregoing sequences, for example, 5-methyl-cytosine replacing cytosine, and (5) nucleotide sequences that are for example, within a 20% and preferably 10% variance to the foregoing nucleotide sequences.

Since nucleotide codons are redundant, also within the scope of this invention are equivalent nucleotide sequences which include: nucleotide sequences which code for or can be translated into the contortrostatin proteins, their protein variants, functional equivalents, or derivatives. These nucleotide sequences may also be used in the practice of the invention.

In addition to the above, contortrostatin nucleotide sequences also include: (1) nucleotide sequences that are capable of hybridizing to the coding sequences of the respective nucleotide sequences, under stringent hybridization conditions, and (2) fragments of or mutagenized nucleotide sequences of those disclosed herein which can encode or can be translated into proteins having substantially the same biological characteristics/activities of the respective contortrostatin proteins, e.g. integrin antagonist, zinc-binding, proteinase, anti-angiogenic factor, and further activities as disclosed herein in further embodiments and examples.

The terms "contortrostatin proteins", as used in relation to proteins include the respective proteins described in the Example section, below, and precursor proteins obtainable by the methods of the present invention, most preferably proteins exhibiting the properties of a disintegrin obtainable from the isolation methods of the Example

15

20

25

30

WO 00/18421 PCT/US99/22608

13

section below, and: (1) protein variants of these proteins; e.g. these protein variants may contain amino acid sequences that have for example, at least 90% or more preferably at least 95% of their amino acids matching the sequences of the pro-protein, metalloproteinase, disintegrin and/or native contortrostatin regions of the protein; (2) the functional equivalents of these proteins and their variants, respectively; and (3) the derivatives, including fragments, of the contortrostatin proteins and their variants, respectively.

The variants can result from, e.g. substitution, insertion, or deletion of the amino acid sequences of the contortrostatin proteins. The derivatives of the proteins and their variants, include fragments of these proteins and immunorective peptides that specifically bind with antibody to contortrostatin.

Two amino acid sequences are functionally equivalent if they have substantially the same biological activities such as the ability to bind to integrin receptors. The proteins may be fused to other proteins, for example, signal sequence fusions may be employed in order to more expeditiously direct the secretion of recombinant contortrostatin proteins.

Substitutional variants of the proteins disclosed herein are those in which at least one residue in the disclosed sequences has been removed and a different residue inserted in its place. Preferably, the amino acid change is conservative. Thus, modifications of the contortrostatin proteins' primary amino acid sequences also include conservative variations. The term "conservative variation" as used herein denotes the replacement of an amino acid residue by another, biologically similar residue. Examples of conservative variations include the substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as the substitution of arginine for lysine, glutamic for aspartic acid, or glutamine for asparagine, and the like. The term "conservative variation" also includes the use of a substituted amino acid in place of an unsubstituted parent amino acid provided that the polypeptide retains its biological activity, e.g., antibodies raised to the substituted polypeptide also immunoreact with the unsubstituted polypeptide.

Further, as is the case for all proteins, the precise chemical structure depends on a number of factors. As ionizable amino and carboxyl groups are present in the molecule, a particular protein may be obtained as an acidic or basic salt, or in neutral form. All such

15

20

WO 00/18421 PCT/US99/22608

14

preparations which retain their activity when placed in suitable environmental conditions are included in the definition. Additionally, the primary amino acid sequence may be augmented by derivatization using sugar moieties (glycosylation) or by other supplementary molecules such as lipids, phosphate, acetyl groups and the like, more commonly by conjugation with saccharides. The primary amino acid structure may also aggregate to form complexes, most frequently dimers. Certain aspects of such augmentation are accomplished through post-translational processing systems of the producing host; other such modifications may be introduced *in vitro*. In any event, such modifications are included in the definition so long as the activity of the protein is not destroyed. It is expected that such modifications may quantitatively or qualitatively affect the activity, either by enhancing or diminishing the activity of the protein in various assays.

Individual amino acid residues in the chain may also be modified by oxidation, reduction, or other derivatization, and the protein may be cleaved to obtain fragments which retain activity. Such alterations which do not destroy activity do not remove the protein sequence from the definition. The following discusses some of the modifications in further detail by way of example.

Thus, glycosylation variants are included within the scope of contortrostatin proteins. They include variants completely lacking in glycosylation (unglycosylated) and variants having at least one less glycosylated site than the native form (deglycosylated) as well as variants in which the glycosylation has been changed.

As illustrated in the examples, native CN may be isolated from the venom of Agkistrodon contortrix contortrix in a relatively straightforward manner. Alternatively, CN may also be prepared by exploiting a variety of biochemical methods in current use, such as recombinant DNA technology or the like. Moreover, the sequence information reported herein can be used for making probes to identify variants, fragments, conserved domains or pro-proteins having substantial homology to CN and its precursor(s). Once identified, the genes may be isolated, further manipulated, and cloned into expression vectors.

There is also provided a vector containing a DNA molecule encoding a contortrostatin protein made according to techniques understood by those with skill in the

10

30

WO 00/18421 PCT/US99/22608

15

art with reference to the disclosure herein. In the present invention, the contortrostatin nucleotide sequence may be inserted into a recombinant expression vector. The term "recombinant expression vector" refers to a plasmid, virus or other vehicle known in the art that has been manipulated by insertion or incorporation of a contortrostatin genetic sequence. Such expression vectors contain a promotor sequence which facilitates the efficient transcription of the inserted genetic sequence in the host. The expression vector typically contains an origin of replication, a promoter, as well as specific genes which allow phenotypic selection of the transformed cells. These vectors may be used to transform competent hosts to produce transformants that are capable of producing the snake venom protein.

Further, there is provided, a prokaryotic or eukaryotic host cell stably transformed or transfected by the vector, as well as a method of making the contortrostatin protein or its biological variants. The method includes the steps of, first culturing a prokaryotic or eukaryotic host cell transformed with DNA encoding for contortrostatin protein; and then, recovering the contortrostatin protein. The host cell can be mammalian cells, plant cells, insect cells, yeast and other fungi, or bacteria.

Sources of DNA sequences encoding for the proteins include isolated DNA from suitable cells or cell lines, cloned DNA from a snake venom genomic library or cloned DNA from a complementary DNA library, where the total complementary DNA is reverse transcribed to DNA and cloned. Once a DNA sequence is identified which encodes for a protein of interest, the sequence of bases may be determined by known means [e.g., Maxam and Gilbert, *Proc. Natl. Acad. Sci. USA* 74:560 (1977)]. In addition, hybrid DNA technology may be employed for obtaining expression. The DNA sequence may be restriction mapped and appropriate sites for cleavage defined. In this way, the sequence may be excised and introduced into a vector having the appropriate regulatory signals. A more detailed discussion of suitable techniques for identification and expression of disintegrin genes is provided in, e.g., U.S. Patent Nos. 5,182,260 and 5,196,403 to Maraganore et al., the entire disclosures of which are hereby incorporated by reference.

Further, the sequence encoding the native protein may then be manipulated (for example, by single or multiple mutations or deletions) in a manner well known in the art to provide modified proteins, in which changes of one or more amino acids have been

20

30

WO 00/18421 PCT/US99/22608

16

introduced. Following the procedures described herein, the determination of whether a particular polypeptide exhibits an activity profile characteristic of CN would then be a matter of routine experimentation. Accordingly, the present invention contemplates both the native CN and mutations thereof which exhibit the characteristic activity profile defined herein. Moreover, CN may also be employed in the form of a fusion protein with a suitable thrombolytic agent. Fusion proteins of this type may be prepared in a manner analogous to that described for formation of platelet aggregation inhibitor/anti-thrombin polypeptide fusion proteins in the aforementioned U.S. Patent Nos. 5,182,260 and 5,196,403 to Maraganore et al.

Despite the high sequence homology with other disintegrins (shown in Fig. 1A) [Niewiarowski, S., McLane, M. A., Kloczewiak, M., and Stewart, G. J. Disintegrins and Other Naturally Occurring Antagonists of Platelet Fibrinogen Receptors, Seminars in Hematology. 31:289-300 (1994)], contortrostatin displays a unique amino-terminal sequence with a truncation of about 8 amino acids. In comparison with other disintegrins, two cysteine residues are missing at the amino-terminus of contortrostatin. It is well established that monomeric disintegrins possess an even number of cysteines, all of which are involved in disulfide bond formation. The missing two cysteines in the contortrostatin precursor may lead to disruption of disulfide bond pairing. This in turn may lead to the formation of two intermolecular disulfide bonds between two identical chains to form the homodimeric structure of contortrostatin. Fig. 5 presents a hypothetical model of the structure of the dimeric disintegrin as compared to a disintegrin, kistrin, whose disulfide bond positioning is known [Adler, M., Carter, P., Lazarus, R. A., and Wagner, G. Cysteine pairing in the glycoprotein IIb/IIIa antagonist kistrin using NMR, chemical analysis, and structure calculations, Biochemistry 32:282-9 (1993)]. In Fig. 5A, the sequences of the two disintegrins are compared. It can be readily seen that kistrin, with 12 half-cystines, forms 6 disulfide bonds. By comparison, contortrostatin, which has two unpartnered cysteine residues, can form the homodimeric structures presented in Fig. 5B. We are not sure at present which alignment is preferred, but we are confident, based on mass spectral analysis and SDS-PAGE of native and reduced contortrostatin, that a homodimeric structure most probably identical to one of the two shown, is formed.

17

CN is clearly quite different from applaggin, as the latter has been demonstrated unmistakably to be a monomer [Wencel-Drake et al. (1993), *supra*]. Moreover, CN does not inhibit platelet release reactions, as has been demonstrated to be the case with applaggin in the aforementioned U.S. Patent Nos. 5,182,260 and 5,196,403 to Maraganore et al. Finally, despite the similarities in sequence there are also significant differences between the sequences with respect to both the start site and the ensuing sequences.

Methods of Use

10

15

20

25

30

CN has been found to be a potent inhibitor of human, rabbit and canine platelet aggregation *in vitro*. Unlike applaggin, however, CN does not inhibit platelet release reactions. Platelets comprise a plurality of different granules, including alpha granules and dense granules, whose contents are released when the platelets aggregate. The finding that CN does not inhibit platelet release of granule contents (including ATP from the dense granules) signifies that the platelets may still release their contents (and thus maintain some semblance of normal physiological activity) notwithstanding the inhibition of aggregation. By contrast, when applaggin inhibits platelet aggregation it also inhibits platelet release (as measured by, e.g., inhibition of serotonin release from the dense granules). Thus, normal platelet physiological processes are necessarily further perturbed with administration of applaggin.

Several lines of evidence indicate that CN inhibits platelet aggregation by binding specifically to the (GP)IIb/IIIa integrin receptor. For example, in a fibrinogen-(GP)IIb/IIIa ELISA [Dennis, M.S. et al., *Proc. Natl. Acad. Sci. (USA)* 87:2471 (1990)], in which the extent of purified (GP)IIb/IIIa bound to immobilized fibrinogen can be quantitated, CN effectively blocks (GP)IIb/IIIa binding. Additionally, the partial amino acid sequence of CN indicates considerable similarity with other disintegrins known to bind to (GP)IIb/IIIa. Finally, CN blocks 7E3 binding to (GP)IIb/IIIa. 7E3 is a murine monoclonal antibody that specifically binds to (GP)IIb/IIIa, thereby inhibiting human and canine platelet aggregation [Coller, B.S. et al., *J. Clin. Invest.* 72:325 (1983)]. In the presence of a low concentration of CN, 7E3 binding to platelets is significantly inhibited.

Three snake venom disintegrins, kistrin [Yasuda et al. (1990), *supra*], echistatin [Holahan et al. (1991), *supra*] and bitistatin [Shebuski, R.J. et al. (1990), *supra*], have

15

20

25

30

WO 00/18421 PCT/US99/22608

18

demonstrated a potential role as antithrombotic agents for use in thrombolytic therapy by enhancing and sustaining arterial thrombolysis in conjunction with recombinant tissue plasminogen activator. Based on the low IC₅₀ values of CN, its *in vivo* efficacy as an antithrombotic agent has been examined. Using a canine reoccluding carotid arterial thrombosis model, CN has been found to efficiently sustain opening of the carotid artery in conjunction with anisoylated plasminogen streptokinase activator complex (APSAC). APSAC alone was found insufficient to prevent the rapid reocclusion of the carotid artery. Heparin was not needed to sustain opening when CN was administered with APSAC. This is a significant distinction over other disintegrins (e.g., echistatin, bitistatin and kistrin) which have been evaluated in models of coronary artery thrombosis.

The compositions of the present invention are particularly useful for treatment of thrombotic diseases in mammals, alone or in conjunction with one or more thrombolytic agents. In particular, the compositions of the present invention have utility in treating or preventing arterial, venous and microvascular thrombosis and thromboembolism. Thus, the compositions have utility in treating stroke, transient ischemic attacks, arteriosclerosis, atherosclerosis, pulmonary embolism, aneurisms and angina. In particular, the compositions have utility in preventing or treating myocardial infarctions.

The compositions of the present invention also have utility in inhibiting metastasis in melanoma, carcinoma and sarcoma patients. CN has been observed to bind to at least two sites on human melanoma M24met cells: a high affinity site with a dissociation constant (Kd) of 1.1 (\pm 0.7) nM and 96,000 (\pm 39,000) sites per cell; and a lower affinity site with a Kd of 41 (\pm 13) nM and 480,000 (\pm 90,000) sites per cell. Moreover, CN has been found to inhibit human melanoma M24met cell adhesion to fibronectin and vitronectin, and to a lesser extent to collagen and laminin. Thus, methods and compositions for preventing metastases in melanoma, carcinoma and sarcoma patients are also contemplated as within the scope of the present invention.

In accordance with the present invention, the unique properties of disintegrins are exploited in methods and compositions for preventing metastasis in carcinoma, sarcoma and melanoma patients. In particular embodiments the positions and methods are provided for preventing metastasis in breast cancer patients.

15

20

25

30

WO 00/18421 PCT/US99/22608

19

In further embodiments, we provide the protein contortrostatin from southern copperhead snake venom that possesses potent anti-tumor activity. A sophisticated technique has been developed to purify this protein from the complex mixture of proteins found in southern copperhead venom. As indicated above, originally CN was characterized as an inhibitor of platelet aggregation. We have purified several disintegrins from snake venoms. Contortrostatin (CN) was purified from southern copperhead venom. Disintegrins contain a constrained Arg-Gly-Asp (RGD) sequence at the tip of a flexible peptide loop, of about 13 amino acid residues flanked by Cys residues, protruding from the main protein core. See, e.g. amino acid residues 457 to 469 of SEQ ID NO: 1 or 2. This exposed RGD sequence enables disintegrins to bind to integrins with high affinity.

We have developed a metastatic breast cancer model by implanting human breast cancer cells into the mammary fat pads of mice. The mice we use were genetically manipulated so that their immune system is deficient and they are unable to reject the implanted human cancer cells. We observed that palpable tumor masses developed in the mammary fat pads two weeks after cancer cell implantation, and that tumor cells spread to the lungs in untreated animals within 12 weeks. CN or placebo was injected daily into tumors in several different groups of mice. Following treatment we found that the size of the tumor masses in the CN treated mice were significantly smaller than those in placebotreated mice. Significantly, the CN-treated group showed >90% inhibition of tumor spread to other sites in the body (metastasis), as compared to the placebo group. Our studies indicate that CN blocks the attachment of breast cancer cells to proteins which are essential components of blood vessel walls. CN also inhibited new blood vessel formation (neovascularization) induced by breast cancer cells following incubation on a chick embryo membranous respiratory organ called the chorioallantoic membrane while placebo treatment did not. Since neovascularization is critical to continued proliferation of a growing tumor the ability to inhibit the growth of new vessels is an important anti-cancer action of CN.

Based on these studies it appears that disintegrins such as the snake venom protein contortrostatin possess anti-metastatic activity. Our findings suggest that CN blocks several critical steps in metastasis and is, therefore, more potent than other agents which only block a single step.

15

20

WO 00/18421 PCT/US99/22608

20

The disintegrin-containing compositions of the present invention are also useful in treatment of osteoporosis. Osteoclasts are multinucleated cells up to 400 µm in diameter which resorb mineralized tissue in vertebrates. Bone resorption appears to proceed by a combination of processes involving attachment to bone, polarized secretion of acid and proteases, and active motility of osteoclasts along the bone substrate; osteoclasts bind to bone via an RGD-sequence as an obligatory step in bone resorption, and this RGD-binding integrin is at adhesion structures [Sato, M. et al., *J. Cell Biol.* 111:1713 (1990)]. The molecular mechanisms whereby osteoclasts attach to bone are not well understood; however, by analogy to other cells, members of the integrin superfamily of divalent cation-dependent adhesion molecules are believed to mediate this interaction. Disintegrins, such as echistatin [Sato et al. (1990), *supra*] and presumably CN, inhibit bone resorption by isolated osteoclasts; the mechanism of action is presumably by disrupting adhesion structures. Accordingly, compositions and methods for treatment of osteoporosis employing an amount of CN effective to inhibit bone resorption by osteoclasts are also contemplated as within the scope of the present invention.

Finally, CN has utility in the promotion of wound healing. Events involved in wound healing are known to include alterations in integrin expression or functional activity and suggest that integrin receptor modulation plays a central role in wound repair and inflammation. Fibronectin is also known to play a number of roles in the wound healing process. Although fibronectin function is thought to be critical to effective wound healing, there are reports that at least one of its activities (the binding of bacteria) may be counterproductive [Grinnell, F., J. Cell. Biochem. 26:107 (1984); Clark, R.A.F., Arch. Dermatol. 124:201 (1988)]; the presence of fibronectin in the wound bed may thus promote bacterial attachment and infection. Fibronectin also appears to be intimately involved in keloid formation. Keloids are a pathological consequence of wound healing that affects a significant proportion of non-Caucasian patients. Keloids are benign tumors of connective tissue that grow beyond the boundary of the original wound and are rich in fibronectin and type I collagen [Sible, J.C. & Oliver, N., J. Cell. Biochem. Suppl. 16F:170 (1992)]. By virtue of their inhibition of cell-cell and cell-extracellular matrix interactions (including interaction with fibronectin), disintegrins such as CN would be expected to have a profound effect on processes involved in wound repair, including keloid formation.

15

20

25

30

WO 00/18421 PCT/US99/22608

21

A major problem following obstetrical and gynecological surgery is the formation of adhesions. This widespread phenomenon observed in peritoneal wound repair is a leading cause of pain, intestinal obstruction and infertility. Adhesion formation appears to involve an imbalance in the fibrinolytic and fibroproliferative inflammatory responses and may also involve a modulation of the cell-cell or cell-extracellular matrix interactions. There is strong evidence for an important role of fibrin during the initial stages of adhesion formation [diZerega, G.S., *Prog. Clin. Biol. Res.* 381:1 (1993)]. The presence of cellular elements, including platelets, further exacerbates the role of fibrin. In view of the role of platelets and fibrin in adhesion formation, the use of disintegrins such as CN as a potential therapeutic agent is most attractive.

In preliminary studies in a rabbit model of adhesion formation, abrasion and devascularization of the uterine horns of rabbits were employed to induce adhesion formation during wound healing in untreated animals [Rodgers, K. et al., *Int. J. Fertil.* 35:40 (1990)]. Alzet pumps were employed to continuously deliver CN at a rate of 10 µl/hr (36 µg/ml). In this model system, decreased adhesion formation was observed in treated animals compared to controls. Therefore, compositions and methods for preventing adhesion formation whereby an amount of CN effective to prevent adhesion formation is administered to a patient in need of such treatment are also contemplated as within the scope of the present invention.

The compositions of the present invention comprise at a minimum an amount of CN effective to achieve the desired effect (i.e., prevent thrombus formation, prevent metastasis in carcinoma patients, prevent adhesion formation, etc.) and a suitable carrier or excipient. Generally, in these compositions, CN is present in an amount sufficient to provide about 0.01 mg/kg to about 50 mg/kg per day, preferably about 0.1 mg/kg to about 5.0 mg/kg per day, and most preferably about 0.1 mg/kg to about 0.5 mg/kg per day. Such compositions have particular utility in the prevention of thrombus formation.

Alternatively, CN is administered in combination with at least one thrombolytic agent present in an amount effective to achieve thrombolysis. Suitable thrombolytic agents include, but are not limited to, the following: anisoylated plasminogen streptokinase activator complex (APSAC); tissue-type plasminogen activator (tPA);

22

urokinase-type plasminogen activator (uPA); and fibrolase, a snake venom fibrinolytic agent as described in U.S. Patent 4,610,879 to Markland, Jr. et al.

CN may be administered by a variety of heretofore known means suitable for delivery thereof into the blood stream in substantial amounts. Intravenous administration of CN in a suitable liquid vehicle or excipient is presently contemplated as the preferred route of administration. CN is soluble in water, and may therefore be effectively administered in a suitable aqueous solution (e.g., phosphate buffered saline). Alternatively, CN may be administered orally (in the form of tablets or capsules formulated with a suitable binder or excipient material, or in the form of aqueous or oily suspensions, solutions, emulsions, syrups or elixirs) or as a parenteral suspension. As is well known in the art, adjuvants such as local anesthetics, preservatives, buffering agents, lubricants, wetting agents, colorants, flavorings, fillers and diluents may suitably be included in any of these formulations.

15 EXAMPLES

10

These additional embodiments may be better understood with reference to the accompanying examples, which are intended for purposes of illustration only and should not in any sense be construed as limiting the scope of the invention as defined in the claims appended hereto.

In conducting the following examples, lyophilized venom from Agkistrodon contortrix contortrix was obtained from Biotoxins, Inc., St. Cloud, FL. All chemicals were of the highest grade available. Pierce protein assay kit using bicinchoninic acid was employed to determine protein concentrations [Smith, P.K. et al., Anal. Biochem. 150:76 (1985)].

For hydrophobic interaction (HIC)-HPLC a Perkin Elmer 410 LC pump was employed with a LC-95 UV/VIS detector. For reverse phase HPLC a Spectra Physics LC 8810 pump was employed with an SP 8450 UV/VIS detector. Absorbance for HIC-HPLC was monitored at 280 nm and for RP-HPLC at 215 nm. A polypropyl aspartamide (250 x 21 mm) column (Poly LC, Columbia, MD) was used for hydrophobic interaction HPLC.

C18 (218TP54 and 218TP510) columns were used for reverse phase (RP) HPLC (Vydac,

23

Hesperia, CA). For cation exchange chromatography a CM (carboxymethyl) 300 column (SynChrom, Inc., Lafayette, IN) was employed.

Example 1

10

15

20

25

30

5 Purification and characterization of CN

CN was purified from Agkistrodon contortrix contortrix (Southern copperhead) venom using a four step HPLC procedure. For the first step of purification crude venom (1 g) was dissolved in 0.1 M phosphate buffer containing 1 M ammonium sulphate, pH 6.8 (buffer A) and applied to the polypropyl aspartamide HIC-HPLC column. Elution was achieved as follows: 50 minutes isocratically with 100% buffer A; a linear gradient for 90 minutes to 0.1 M phosphate, pH 6.8 (buffer B); 40 minutes isocratic at 100% buffer B. Fractions of 10 ml were collected in a Pharmacia Frac 100 fraction collector at 4° C using a flow rate of 5 ml/min. Fractions containing platelet aggregation inhibiting activity were pooled and concentrated by ultra-filtration using an Amicon stir cell with a YM3 membrane. Proteins were detected at 280 nm; platelet aggregation inhibiting activity was observed in the flow through.

Further purification was achieved by C18 RP-HPLC. Fractions containing platelet aggregation inhibiting activity were pooled and concentrated for this second step. The C18 column (218TP510) was equilibrated with 95% of 0.1% TFA in water (solvent A) and 5% of 80% acetonitrile in 0.1% TFA in water (solvent B). Elution was achieved as follows: isocratic at 95% solvent A and 5% solvent B for 10 minutes; a linear gradient to 40% solvent B in 65 minutes; linear gradient to 100% solvent B in 20 minutes; isocratic at 100% solvent B for 25 minutes. Fractions were collected manually every minute at a flow rate of 7 ml/minute. CN eluted at 28% acetonitrile (66 minutes).

Fractions containing platelet aggregation inhibiting activity were pooled and rerun on the same C18 RP-HPLC column using a shallower gradient. Elution was achieved as follows: isocratic at 80% solvent A and 20% solvent B for 20 minutes; a linear gradient to 30% solvent B over 90 minutes; and a 25 minute linear gradient to 100% solvent B. CN eluted as a sharp peak at 22% acetonitrile (82 minutes). The minor peak eluting just before CN also contained platelet aggregation inhibiting activity and had a similar

15

20

25

WO 00/18421 PCT/US99/22608

24

molecular weight to that of CN; due to the low yield, this peak was not further characterized.

A final purification step was performed using pooled fractions from the previous step. These pooled fractions were applied to a cation exchange, CM300, HPLC column and elution was achieved by an increasing gradient of sodium chloride. CN elutes at 52.5 minutes (160 mM NaCl). This step achieved a separation of CN from isoforms thereof. Yields of 1-2 mg of the four-step purified CN were obtained per gram of crude venom.

For SDS-polyacrylamide gel electrophoresis (SDS-PAGE) Tris-Tricine 16.5% gel was used according to published protocols under reducing and non-reducing conditions [Schagger, H. & Von Jagow, G., *Anal. Biochem.* 166:368 (1987)]. The gel was run using a BioRad minigel system and stained with silver [Morrisey, J.H., *Anal. Biochem.* 117:307 (1981)] or Coomassie blue R250.

SDS-PAGE analysis of CN revealed that it has a molecular mass of approximately 15,000 Daltons under non-reducing conditions and 5,000-7,000 Daltons under reducing conditions. This strongly suggests that CN is composed of two subunits. Another possibility, albeit unlikely, is that the large difference in migration may be attributed to differential uptake of SDS under non-reducing and reducing conditions.

The molecular weight of CN was confirmed by mass spectrometry using a triple quadrupole instrument with an electrospray ion source. A mass of 13,507 Daltons was determined for intact CN; the analysis also indicated a high degree of purity. Mass spectrometry of the reduced and pyridylethylated protein gave a mass of 7,996 Daltons. This is the expected value for the individual chains of a homodimer of this molecular weight, taking into account the incorporation of 1,248 mass units for the 12 pyridylethyl groups incorporated into the 6 reduced disulfide bounds (based on homology with known disintegrins, there should be 6 disulfide bonds). These findings place CN in a unique position among all the disintegrins reported to date in that it exists as a dimer. Scatchard analysis of CN binding to unactivated human platelets revealed a single class of binding sites with a dissociation constant (K_d) of 37 nM and number of binding sites (B_m) equal to 100,000. Reduction of the disulfide bonds completely eliminated platelet aggregation inhibitory activity, even at concentrations ten times the IC₅₀, suggesting that structural parameters are critical for maintaining activity.

25

Example 2

20

30

cDNA Cloning of Contortrostatin Using Polymerase Chain Reaction

Partial amino acid sequence analysis of contortrostatin using Edman degradation 5 methods suggested that the subunit of contortrostatin is homologous with other disintegrins (Fig. 1A) with the cysteine residues aligned, as well as the RGD sequences [Niewiarowski, S., McLane, M. A., Kloczewiak, M., and Stewart, G. J. Disintegrins and Other Naturally Occurring Antagonists of Platelet Fibrinogen Receptors, Seminars in Hematology 31: 289-300 (1994)]. The strategy for cloning contortrostatin cDNA with 10 PCR is based on the structural homology among the disintegrin family. Design of the PCR primers is schematically illustrated in Fig 1. The underlined sequences in Fig. 1A are highly conserved among disintegrin family. PCR primers were synthesized based on this region. The nucleotide sequence coding this region in the cDNA of trigramin from Trimeresurus gramineus was used to synthesize the primers PCR-1 and PCR-2 (Fig. 1B). PCR-1 and PCR-2 are complimentary primers corresponding to the coding sequence of a consensus sequence PCCDAATCKL among disintegrins. Their nucleotide sequences are:

PCR-1: 5'-GTTTACAGGTTGCAGCATCGC-3' (SEQ ID NO: 3)

PCR-2: 5'-GCGATGCTGCAACCTGTAAAC-3' (SEQ ID NO: 4)

λgt 10 forward and reverse primers which flanks the EcoRI site of the vector were used for PCR. Nucleotide sequences of the primers are listed below:

25 \(\lambda\text{gt10 forward primer: 5'-AGCAAGTTCAGCCTGGTTAAG-3' (SEQ ID NO:5)}\) \(\lambda\text{gt 10 reverse primer: 5'-CTTATGAGTATTTCTTCCAGGGTA-3' (SEQ ID NO: 6)}\)

Oligonucleotide primers were synthesized by the Microchemical Core Facility of the University of Southern California Comprehensive Cancer Center. Primers were provided in a deprotected lyophilized form, and were resuspended and diluted to the appropriate concentration with water prior to use.

26

Example 3

5

10

15

20

25

30

PCR amplification of contortrostatin cDNA

We used the cDNA library of *Agkistrodon contortrix contortrix* venom gland constructed in λgt 10 vector at the EcoRI site. The estimated titer of the library was 10¹⁰ plaque-forming units (pfu)/ml and the complexity was 50,000. 500μl of the cDNA library phage solution was mixed with 500μl of 20% polyethylene glycol (PEG)/ 1M NaCl solution in an Eppendorf tube. The Eppendorf tube was inverted twice and incubated at room temperature for 30 min. The solution was then centrifuged at 14,000 rpm for 10 min. The supernatant was discarded and the pellet was resuspended in 100 μl of sterile water. The suspension was incubated with 10μl of proteinase K (10 mg/ml) at 50°C for 1 hr. The phage particle suspension was extracted with phenol/chloroform twice and the DNA was precipitated with 1/10 volume of 3M sodium acetate and 2 volumes of absolute ethanol followed by washing with 80% ethanol. DNA was resuspended in 10μl of sterile water in preparation for PCR.

The PCR reaction was set up as follows: 5µl of DNA solution was mixed with 1µl of 25 mM dNTP's (Pharmacia), 1 µl (100 ng/µl) each of forward and reverse PCR primers, and 5 µl of 10x PCR buffer. The final volume of the reaction was brought up to 50 µl with water. After mixing, one drop of mineral oil was applied on top of the liquid. The Eppendorf tubes were pre-heated to 98°C for 5 min, then incubated at 70°C and 60°C for 1 min respectively before cooled down to room temperature. 2.5 units of Taq DNA polymerase (Pharmacia) was added to each mix. The thermal cycler was programmed as follows: 96°C for 15 sec, 55°C for 30 sec, 72°C for 1 min. PCR amplification was performed for 30 cycles. Following the last cycle, a final 72°C extension step for 7 min was allowed before the samples were cooled down to 30°C. The PCR product was extracted with chloroform before it was analyzed by agarose gel electrophoresis. The bands resolved by electrophoresis were recovered from the agarose gel using Geneclean kit (Bio 101, Inc.) according to the manufacturer's manual for direct DNA sequencing.

λgt10 forward and PCR-1 primers were used to amplify the up stream region of the adhesion site (Fig 1B). Similarly, PCR-2 and λgt10 reverse primers were coupled to amplify the down stream part of the cDNA (Fig. 1B). A major band of about 1300 bp (designated CN-N) was obtained with the first pair of primers (Fig. 2, lane 1), and a major

27

band of approximately 700 bp (designated CN-C) resulted from PCR using the later pair of primers (Fig. 2, lane 2). CN-N and CN-C were subjected to nucleotide sequencing analysis prior to the overlapping extension. As expected, CN-N demonstrated high similarity to cDNA of trigramin which encodes its N-terminal signal peptide. CN-C nucleotide deduced amino acid sequence was very similar to the disintegrin COOH-terminal sequence encoding the RGD site.

Since PCR-1 and PCR-2 were complimentary, CN-N and CN-C overlap at this site, and therefore, can be assembled into a full length cDNA. To accomplish this goal, we used an overlapping extension method as shown in Fig. 1C. Briefly, equal molar amount of double-strand PCR products CN-N and CN-C were mixed with λ gt10 forward and reverse primers. After denaturation of both double-strands, the subsequent reannealing results in two kinds of molecules. One is annealed CN-N and CN-C at the adhesion site with the recessive ends as 3'-ends. This molecule can be automatically elongated into full length double-strands using PCR. The other molecule is similarly annealed, but with the recessive ends as 5'-ends. Although these molecules are not self-elongated, the recessive parts can be filled in with priming by λ gt10 primers at each end. Fig. 2, lane 3 shows the products of overlapping extension as resolved by agarose gel electrophoresis. The size of the major band is estimated to be 2,000 bp which equals the sum of CN-N and CN-C, and is thus designated "full length". The full length band was recovered from the gel and treated with EcoRI. Subsequently, this piece of DNA was subcloned into the plasmid vector pcDNA3.1(+).

Example 4

10

20

25

Subcloning of PCR product into plasmid vector

Plasmid pcDNA3.1(+) was digested with EcoRI (Pharmacia) followed by dephosphorylation using T4 phosphatase (Boehringer-Mannheim). PCR overlapping extension product was also digested with EcoRI. The PCR product was inserted into the linearized vector by ligation reaction using T4 ligase (Pharmacia) at 16°C overnight. All the reactions were set up and performed according to standard protocols. Successful ligation was selected by plating transformed *E. coli* (DH5α) on ampicillin containing

28

plates. Plasmids containing the insert were amplified in *E. coli*. Purified plasmid DNA was obtained with Qiagene DNA Miniprep columns.

Example 5

5 cDNA sequencing

Automated DNA sequencing was performed by the Microchemical Core Facility. PCR primers were used as sequencing primers for direct sequencing of PCR products. For sequence analysis of the insert in plasmid pcDNA3.1(+), T7 promoter primer and BGH reverse primer, which flanks the multiple cloning sequence (MCS), was utilized to initiate the assay. Typical reactions gave readable sequences of 400 to 600 bp. The sequencing reactions were performed on double-stranded DNA in the case of plasmid DNA. With the synthesis of new sequencing primers, additional sequences were obtained, and these were assembled into overlapping contiguous sequences using DNAsis computer program.

Fig. 3 shows the full length nucleotide sequence inserted between the EcoRI sites. It is composed of 2,029 nucleotides, which is the size of the full length band of the overlapping extension (Fig. 2, lane 3). Following an 86-nucleotides 5'-end non-translated region (5'-NTR), an open reading frame is found between nucleotides number 87 and 1535. Nucleotide 1536 to 1538 is the termination codon. The 3'-NRT possesses an AATAAA site in the 3'-end noncoding region, and ends with poly(A) tail, suggesting that 20 the cDNA we obtained with the overlapping extension was indeed a complete cDNA (Fig. 3). The open reading frame encodes 483 amino acids. The structure of cDNA deduced amino acid sequence can be divided into three domains. The first 190 amino acids, starts with methionine, are highly similar to the pro-protein of many cloned snake venom proteins (comparison is shown in Fig. 4). From amino acid 191 to 418 is the 25 metalloproteinase domain including a Zinc-binding motif HEMGHNLGISH (aa 334 to 344). The remaining 65 amino acids belong to contortrostatin monomer which is identical to the known partial amino acid sequence of contortrostatin determined by Edman methodology. This sequence is very similar to those of many disintegrins whose sequences have been determined (Fig. 4 and 1A). The calculated molecular weight of the 30 disintegrin is 6.77 kDa, which is equal to that of CN monomer. The RGD (aa 461 to 463) sequence is in bold face letters. The three-domain structure matches the precursor model

29

of snake venom metalloproteinase and disintegrin proposed by Kini et al. [Structural Domains in Venom Proteins: Evidence that Metalloproteinases and Nonenzymatic Platelet Aggregation Inhibitors (Disintegrins) from Snake Venoms are Derived by Proteolysis from a Common Precursor, *Toxicon* 30:265-296, (1992)]. There is evidence that disintegrins are synthesized in the snake venom gland cells as a multi-domain precursor which undergo post-translational proteolysis and folding to generate a mature disintegrin.

Example 6

10

15

20

Assay of platelet aggregation inhibitory activity

Column fractions obtained during purification were assayed for activity using fresh human platelet rich plasma (PRP) prepared from blood obtained from human volunteers who had had no medication for at least two weeks. Blood (36 ml) was drawn into 4 ml of 0.1 M citrate and centrifuged at 150 x g for 20 minutes. The supernatant, PRP, was removed and the remaining blood was centrifuged at 10,000 RPM to obtain platelet poor plasma (PPP). Platelet counts were adjusted to 250,000 platelets/µl using a Coulter counter. A Helena four channel aggregometer was used to monitor platelet aggregation. Inhibition of ADP-induced platelet aggregation was monitored at 37° C by adding venom fractions one minute prior to the addition of ADP (10-20 µM final concentration). Fractions exhibiting platelet aggregation inhibiting activity were pooled and further purified. Rabbit and canine PRP was prepared by the same procedure and used in the studies described below.

CN inhibited ADP-induced platelet aggregation in human, canine, and rabbit PRPs (Fig. 6). Empty circles represent human platelet rich plasma, solid circles represent canine PRP, and empty triangles represent rabbit PRP. Varying concentration of CN were preincubated for one minute with PRP prior to the addition of ADP. CN (0.73 μg/ml) inhibited 10 μM ADP-induced human platelet aggregation by 50% (IC₅₀). The IC₅₀ for 20 μM ADP-induced canine platelet aggregation was 1.8 μg/ml for CN. Interestingly, the IC₅₀ for CN mediated inhibition of rabbit platelet aggregation was considerably higher; the IC₅₀ for 20 μM ADP-induced rabbit platelet aggregation was 17.3 μg/ml for CN.

30

Example 7

30

Measurement of (GP) IIb/IIIa specific binding

10

15

25

Measurement of CN binding to platelet (GP)IIb/IIIa receptor was carried out using PRP prepared from blood obtained from human volunteers or male mongrel dogs. PRP was prepared as described above and the platelet count was determined with a H-10 cell counter (Texas International Laboratories, Inc., Houston, TX). PRP (180 µl) was incubated with 20 µl of varying concentrations of CN at room temperature. Radiolabelled antibody (125I-7E3 IgG, 20 µl, 18 mg/ml, 80,000 cpm), specific for (GP)IIb/IIIa, was then added and the mixture incubated for 30 minutes. To establish equilibrium binding, 50 µl aliquots of the binding assay mixture were layered over 200 µl of 30% sucrose in 0.4 ml microcentrifuge tubes and spun at 10,000 RPM for 4 minutes in a swinging bucket rotor to separate platelet-bound antibody from free antibody. The pellet and the supernatant were separated and counted in a Packard Minaxi 5000 series gamma counter. The number of molecules of 125I-7E3 bound per platelet in the presence and absence of CN was calculated by using the following formula:

(4) x $0.9 \mu g$ 7E3 x 3.76×10^{12} molecules 7E3/ μg

(5)

wherein (1) = Pellet counts; (2) = Supernatant counts; (3) = Total CPM (1) + (2); (4) = 20 Fraction bound (1)/(3); and (5) = Platelet counts per μ l x 45 μ l.

The competitive binding studies using 7E3 demonstrated specific platelet (GP)IIb/IIIa receptor binding for CN with both human (Fig. 7A) and canine (Fig. 7B) platelets. The concentration of CN to inhibit 50% of 7E3 binding to human (GP)IIb/IIIa (IC₅₀) is 0.4 μ g/ml. The IC₅₀ for CN for canine (GP)IIb/IIIa is 0.24 μ g/ml. These studies confirm that CN inhibits platelet aggregation by binding to (GP)IIb/IIIa.

31

Example 8

15

20

25

In Vivo Thrombolytic Efficacy of CN

CN has been studied in a reoccluding canine model of arterial thrombosis. The protein was studied initially by systemic infusion at different dosages to determine its relative potency. This data has permitted an assessment of the systemic dose needed for effective antithrombotic (antiplatelet) activity. The effects upon physiological parameters and circulating coagulation factors have also been monitored.

The model of carotid artery thrombosis in the anesthetized canine described is a modification of one developed for the study of experimentally-induced coronary artery thrombosis [Romson, J.L. et al., Thromb. Res. 17:841 (1980)]. The experimental procedure results in the formation of a platelet rich intravascular thrombus at the site of an electrolytically-induced endothelial lesion in proximity to a distal arterial stenosis. The carotid artery is selected for the experimental model, thereby allowing one vessel to be used as a control and the other to be used after administration of the thrombolytic and antithrombotic therapy. APSAC (anisoylated plasminogen streptokinase activator complex) has been used as the thrombolytic agent successfully in this model. The carotid artery response to the electrolytic injury is similar to that observed in the canine coronary artery but has the advantage of each dog demonstrating the ability to form bilateral occlusive thrombi. The lytic-antithrombotic combination of agents may then be administered to only one of the occluded vessels; this allows for an internal control and eliminates those animals that may not form thrombi due to causes unrelated to the vessel wall injury and subsequent occlusive thrombus formation, i.e., low circulating platelet counts, enhanced spontaneous thrombolysis, presence of heart worms, etc. Parameters which are recorded include repeated measures of: phasic and mean carotid artery blood flow velocity using an ultrasonic flow probe, time to thrombotic occlusion, time to recannulization, ex vivo platelet aggregation, prothrombin time, thrombin time, activated partial thromboplastin time, red cell and white cell counts, hematocrit, EKG profile and body temperature, before and after administration of APSAC or APSAC plus CN to separate groups of animals.

Conditioned male mongrel dogs (8-10 kg) have been used for all *in vivo* studies.

Dogs are anesthetized with sodium pentobarbital, intubated and allowed to breath room

15

20

25

30

WO 00/18421 PCT/US99/22608

32

air under positive pressure respiration. Arterial blood gasses and pH determinations are made every 45 minutes and appropriate adjustments made to maintain the blood gasses and arterial pH within normal limits. Both common carotid arteries and the right internal jugular vein are exposed. A catheter is inserted into the jugular vein for blood sampling and administration of the test drug. Arterial blood pressure is monitored from the cannulated femoral artery with the use of a blood pressure transducer. A Doppler flow probe is placed on each common carotid artery proximal to both the point of insertion of the intraarterial electrode and the mechanical constrictor. The constrictor is adjusted until the pulsatile flow pattern is reduced by 50% without altering mean blood flow. Blood flow velocity in the carotid vessels is monitored continuously. Fig. 8 is a schematic representation of the instrumentation of the carotid artery.

Electrolytic injury to the intimal surface of each carotid vessel is accomplished with the use of an intravascular electrode. Each intraarterial electrode is connected to the positive pole (anode) of a dual channel stimulator. The cathode is connected to a distant subcutaneous site. The current delivered to each vessel is monitored continuously and maintained at 300 µA. The anodal electrode is positioned to have the uninsulated portion in intimate contact with the endothelial surface of the vessel. Proper positioning of the electrodes in each of the carotid arteries is confirmed by visual inspection at the end of each experiment. The anodal current is applied for a maximum period of 3 hours or is terminated 30 minutes after blood flow in the involved vessel remains stable at zero flow velocity to verify having achieved formation of a stable occlusive thrombus. The right carotid artery serves as the control vessel, whereas the left carotid artery serves as the test vessel. Vessel wall injury is induced simultaneously in each carotid artery.

APSAC (0.05 U/kg) is infused as a bolus proximal to the thrombus in the left carotid artery only. The dose of APSAC has been determined as one that will consistently lyse the locally injected carotid thrombus without producing a systemic lytic effect. Thus, lysis in the uninjected right carotid should not occur. CN is given intravenously in a 10% bolus immediately following APSAC and the remaining 90% is infused over 1 hour. CN dosages ranged from 0.155 to 0.40 mg/kg; the agent was dissolved at the appropriate dose in a volume of 20 ml of sterile saline for infusion. Reperfusion is defined as the restoration of carotid artery blood flow velocity to 20% of baseline values. Patency is

20

25

30

WO 00/18421 PCT/US99/22608

33

defined as measurable carotid artery flow velocity. Blood pressure, heart rate, and carotid artery flow velocity are monitored for 6 hours or until rethrombosis occurs.

Blood (20 ml) was withdrawn for platelet studies from the jugular cannula into a plastic syringe containing 3.2% sodium citrate as anticoagulant (1/10 citrate/blood, vol/vol). Blood was taken for platelet aggregation and whole blood cell counts at baseline 60, 120, 180, 240 and 300 minutes after the administration of CN. The platelet count was determined with a cell counter. Platelet rich plasma, the supernate present after centrifugation of anticoagulated whole blood at 140 x g for 5 minutes, was diluted with platelet poor plasma to achieve a platelet count of 200,000/mm³. Platelet poor plasma was prepared after the platelet rich plasma was removed, by centrifuging the remaining blood at 12,000 x g for 10 minutes and discarding the bottom cellular layer. Ex vivo platelet aggregation was measured by established spectrophotometric methods with a four channel aggregometer by recording the increase in light transmission through a stirred suspension of platelet rich plasma maintained at 37° C. Aggregation was induced with arachidonic acid (0.65 mM and 0.325 mM) and ADP (20 µM and 5 µM). A subaggregatory dose of adrenaline (550 nM) was used to prime the platelets before stimulation. Values are expressed as percentage of aggregation, representing the percentage of light transmission standardized to platelet rich and platelet poor plasma samples yielding 0% and 100% of light transmission, respectively.

At the conclusion of the study protocol each vessel segment is ligated, proximal and distal to the point of injury, and removed without disturbing the intravascular thrombus. The vessel segment is opened and the intact thrombus is lifted off and weighed.

Five animals have been studied thus far with CN plus APSAC, six with APSAC alone, and a positive control group of six dogs with APSAC plus 7E3 anti-(GP)IIb/IIIa monoclonal antibody. There were essentially no changes in mean arterial blood pressure or mean heart rate following infusion of CN. Further, the carotid artery flow velocity stayed at a high level following infusion of APSAC plus CN as compared to APSAC infusion alone. In the group of animals infused with APSAC alone, the carotid artery opened for a few minutes following infusion of the lytic agent but then reoccluded and remained closed for the duration of the study. In the positive control group, the animals were infused with APSAC (0.1 U/kg) intraarterially and this was followed by a bolus of

34

7E3 anti-(GP)IIb/IIIa F(ab')2 (0.8 mg/kg). In these three animals, the carotid artery remained open following infusion of the combination of APSAC and 7E3 and remained open until the conclusion of the experimental protocol. In the group of five animals infused with APSAC plus CN the results were essentially the same as with the combination of 7E3 and APSAC. However, Table 1 reveals that there was a significant advantage to the combination of APSAC plus CN in terms of the residual thrombus weight. In Table 1, CTTX = CN and RCA = right carotid artery. In the group of five animals treated with this combination of agents the residual thrombus weight per kg dog weight was 1.5, versus 2.4 in the six animals in the APSAC plus 7E3 group, and 4.1 in the APSAC alone group (six animals). Finally, in one of the dogs treated with APSAC plus CN (0.155 mg/kg) platelet aggregation and platelet counts were followed (Fig. 9); CN infusion began at time 0 and was continued for 60 minutes thereafter.

35

TABLE 1

WEIGHT OF RESIDUAL THROMBUS IN CANINE CAROTID ARTERY THROMBOSIS MODEL

5

25

		APSAC Control		APSAC	& 7E3	APSAC	& CTTX
	Dog#	Dog <u>Weight</u> (kg)	RCA <u>Thrombus</u> (mg)	Dog <u>Weight</u> (kg)	RCA Thrombus (mg)	Dog <u>Weight</u> (kg)	RCA Thromb us (mg)
10	1	19.2	27.3	9.4	37.3	15.2	16.8
	2	6.5	47.6	12,2	37.5	8.2	19.2
	3	18.2	90.5	17	53.2	9.2	3.5
	4	20.2	60.1	11.6	3.2	8.5	33.1
	5	16.8	67.5	8.6	36.5	9	15.1
15	6	10	78.5	15.9	11.8		-400
-	Mean	15.2	61.9	12.5	29.9	10	14.6
	SEM	2.1	8.4	1.4	7.6	1.4	4.9
20	Thrombus weight						
	per kg		4.1		2.4		1.5

These results are typical of those in dogs in this group. It can be seen that platelet aggregation was compromised by treatment with the venom protein, but that there appeared to be a return of aggregation at the conclusion of the experiment. Similarly, the platelet count was also depressed during the course of the experiment. It is suspected that the platelets are sequestered in some sanctuary such as the spleen and are then released following a short residence time; platelets returning into the circulation appear to be functional. There is a drop in platelet counts to 10-20% of the baseline value, with

36

aggregability fluctuates somewhat due to the low platelet count; however, it can be seen that platelet aggregability in the residual platelets appears to be returning to normal at the conclusion of the experimental procedure.

5 Example 9

10

15

20

25

30

CN effects on mammary carcinoma adhesion & invasion

The effect of CN on binding of highly metastatic human breast cancer cells, MDA-MB-435 cell line, to ECM proteins was examined. Human fibronectin and vitronectin were immobilized in the wells of 96-well microtiter plates. Referring to Figs. 3 and 4, CN inhibited adhesion of MDA-MB-435 to both ECM proteins in a dose dependent manner. IC₅₀ for adhesion to fibronectin is 18 nM (Fig. 10) and for vitronectin the IC₅₀ is 1.5 nM (Fig. 11). CN had minimal effect on the weak adhesion seen by MDA-MB-435 cells to human type I collagen, or to rat type I collagen to which the MDA cells have a relatively strong affinity. In a variation of the above experiments, CN was immobilized. It was found that CN can support binding of MDA-MB-435 cells in a dose dependent manner. Binding of MDA-MB-435 cells to immobilized CN is blocked by an RGD peptide, GRGDSP (IC₅₀=0.4 mM), and by EDTA (IC₅₀=0.8 mM). Since integrin receptors require metal ions for non covalent association of their subunits, our findings indicate that CN binds to integrin receptors on the surface of MDA-MR-435 cells via an RGD-mediated mechanism. The finding that immobilized CN can support adhesion of MDA-MB-435 cells suggests that this binding involves cell surface receptors on the tumor cells. Referring to Figs. 12 and 13, varying concentrations of GRGDSP (Fig. 12) or EDTA (Fig. 13) were used to inhibit binding of human mammary carcinoma cells to immobilized CN. CN was at 0.1 µg/well. The vertical line at each data point indicates the y-axis error bar. All experiments were conducted as three sets of triplicates for each data point.

Since adhesion of MDA-MB-435 cells to immobilized CN is completely blocked by GRGDSP and by EDTA (Figs. 12 and 13), CN binds solely to integrin receptors of MDA-MB-435 cells via an RGD-dependent mechanism.

Referring to Fig. 14, we have also demonstrated the inhibitory effect of CN on the invasion of a synthetic basement membrane by the MDA-MB-435 cells using a Matrigel-coated invasion chamber. 2.5 x 10³ MDA-MB-435 cells treated with various

37

concentration of CN were allowed to migrate across the Matrigel layer for 48 hrs. Assays at each CN concentration were performed in triplicate. Cells invaded through the Matrigel filter were fixed and stained. Invaded cells were quantitated with microscope by the mean of cell numbers in three randomly selected vision field.

5

10

20

Example 10

CN inhibits growth and metastasis of MDA-MB-435 breast cancer in nude mouse experimental model

A spontaneous (orthotopic) metastatic model of nude mice has been established by implantation of MDA-MB-435 cells (5 x 10⁵ in 0.1 ml) in the mammary fat pads (mfp). Palpable tumors appeared by the 10th day post-implantation. Daily injections of CN into the tumor masses of each of the groups are carried out, started at the 14th day postimplantation. By the 8th week post-implantation, tumors were removed. The animals were allowed to survive for 2 more weeks without CN administration. The animals were then sacrificed and lung metastases were carefully examined. Referring to Fig. 15, our findings indicate that local injection of CN substantially inhibited the growth rate of the tumor. The volumes of tumor masses (mean ± S.D.) of control (dark bars), low-dosage (0.5 μg/day, gray bars), and high-dosage (5μ/day, light gray bars) CN-treated group are shown. The seven clusters of bars from left to right represent the data of pre-injection (PI. 14th day post-implantation) and the 1st through 6th week of injection. Student t-tests were employed to test the significance of differences. * and ** indicate P<0.05 and p<0.01, respectively. The mean weight of tumors treated by high-dose CN (5µg/day) is significantly lower than control group (P<0.05). Table I shows the incidence of lung metastasis based on gross examination and counting of surface nodules. Metastatic spread 25 in the control group is much more extensive than the high dose CN group which showed >90% inhibition of metastasis. These data demonstrate the potential therapeutic role of CN in the treatment of human breast cancer.

38

TABLE 2. Effect of Contortrostatin on the Incidence of Metastatic MDA-MB-435 Breast Cancer in Nude Mouse Experimental Model.

	M	letastasis Incide	nce	
Groups	In situ relapse	Mean size of relapse tumor (mm³)	# of nodules in lung (median)	Other Organs ¹
Control	4/5	66.7 ± 51.7	47.5 ²	5/5
CN (5 μg/day)	2/6	48.7 ± 10.5	4.5	2/6

¹⁰ Organs include: Chest wall, mediastinum, diaphragm, and pleurae

Example 11

5

15

20

25

CN inhibits angiogenesis Induced by MDA-MB-435 tumor In CAM

The hypothesis that inhibitory effect on the growth of the tumors probably results at least in part from the blockage of angiogenesis by CN has been preliminarily verified by observing the effect of CN on tumor induced angiogenesis on chick embryo chorioallantoic membrane (CAM). MDA-MB-435 tumor masses were inoculated on CAM of 10-day chick embryos. CN at various dosages were injected intravenously into CAM on day 2 post-inoculation. Tumor induced angiogenesis and inhibitory effect of CN on angiogenesis can be easily observed in CAM after 3 days of incubation. As shown in Fig. 16, vessels are distributed in a convergent manner with the tumor mass in the center in control embryo. Chick embryo is immunodeficient, and thus allows the growth of implanted MDA-MB-435 tumor. The embryos are incubated at 37°C with humidity at 60%. CN at various dosages was injected intravenously into CAM on day 2 post-inoculation. Tumor induced angiogenesis on the 3rd day is demonstrated the photos. On the top (Fig. 16A) is the control embryo. The vessels are distributed in a convergent manner with the tumor mass in the center. In the middle (Fig. 16B) is the CAM treated with 20 µg of CN. On the bottom (Fig. 16C) is the CAM treated by 150 µg of CN. The

² In 2/5 animals, lungs are directly invaded by cancer cells from pleurae and mediastinum.

39

 $20~\mu g$ CN treated embryo vessels are thinner and less dense than control; tumor mass is smaller than that on control CAM. On CAM treated by $150~\mu g$ of CN the vessels are even thinner and the convergent distribution pattern disappears completely; there is a necrotic tumor mass with volume significantly smaller than control and low dose CN, presumably due to the lack of blood supply (Fig. 16).

Example 12

5

CN has no effect on growth of MDA-MB-435 cells In vitro

MDA-MB-435 cells (0.3 x 106/m]) were added to each well of a 6-well cell culture plates coated with 1/100 dilution of Matrigel. Cells were then treated with CN at various concentration. Growth curves of MDA-MB-435 cells *in vitro* without CN (circles), and with CN at 100 nM (triangles), and 500 nM (diamonds) are illustrated. Sixwell cell culture plates coated with Matrigel (1/100 dilute) were seeded with 3 ml of MDA-MB-435 cell suspension (0.3 x 106/ml). Cell density was determined every 24 hours. Referring to Fig. 17, cells in the presence of CN proliferate equally well as control cells. The result indicate that CN has no direct cytotoxicity during *in vitro* culture of MDA-MB-435 cells.

Example 13

30

20 CN Is effective and well tolerated in vivo

It can be concluded from the chronic experiment with nude mice mentioned above that CN is not toxic. Despite its platelet aggregation inhibitory activity, no spontaneous hemorrhage is observed during the experiment. Some bleeding at the injection sites in CN treated animals, however, was noticed.

25 CN is a novel antimetastatic agent. We hypothesize that CN blocks several critical steps (e.g. adhesion, invasion, angiogenesis) in cancer metastasis and progression. Therefore, it is more potent than other agents which block a single step.

From the foregoing description, one skilled in the art can readily ascertain the essential characteristics of the invention and, without departing from the spirit and scope thereof, can adapt the invention to various usages and conditions. Changes in form and substitution of equivalents are contemplated as circumstances may suggest or render

PCT/US99/22608

40

expedient, and although specific terms have been employed herein, they are intended in a descriptive sense and not for purposes of limitation.

41

What is claimed is:

- 1. A protein consisting essentially of purified contortrostatin, contortrostatin precursor or biologically active variants thereof.
- 2. A recombinant protein according to claim 1.
- 3. A protein according to claim 1 having an amino acid sequence selected from the group consisting of:
 - a) amino acid numbers 419 to 483 of SEQ ID NO: 2;
 - b) amino acid numbers 191 to 410 of SEQ ID NO: 2;
 - c) amino acid numbers 1 to 190 of SEQ ID NO: 2; and
 - d) SEQ ID NO: 2.

5

5

- 4. A purified protein according to claim 1, wherein said contortrostatin comprises a monomer having a molecular mass of about 5 to about 7 kDa.
- 5. A purified protein according to claim 4, wherein the contortrostatin monomer forms a homodimer with another contortrostatin monomer.
- 6. A purified contortrostatin protein according to claim 1 comprising a constrained Arg-Gly-Asp (RGD) sequence at the tip of a flexible peptide loop of about 13 amino acid residues flanked by two Cys residues, wherein the peptide loop is an integrin antagonist which has an amino acid sequence comprising amino acid numbers 457 to 469 of SEQ ID NO: 2.

42

- 7. A protein according to claim 1 wherein said biologically active variants have an amino acid sequence that is at least 90% homologous to:
 - a) amino acid numbers 419 to 483 of SEQ ID NO: 2;
 - b) amino acid numbers 191 to 410 of SEQ ID NO: 2;
 - c) amino acid numbers 1 to 190 of SEQ ID NO: 2; and
 - d) SEQ ID NO: 2.

5

- 8. A peptide recognized by an antibody to contortrostatin.
- A preparation of contortrostatin protein encoded by the nucleotide sequence of SEQ ID NO: 1.
- 10. A preparation of contortrostatin protein according the claim 9, said preparation being substantially free of other snake venom components, wherein said contortrostatin is made by the process of transcribing and translating a cDNA molecule having the nucleotide sequence of SEQ ID NO:1.
- 11. The preparation of claim 10, wherein said transcribing and translating are performed in a recombinant-DNA containing host cell.
- 12. A preparation of contortrostatin substantially free of other snake venom components, wherein said contortrostatin is made by a process of synthesizing a polypeptide having an amino acid sequence according to SEQ ID NO:2.
- 13. A recombinant DNA molecule comprising a DNA sequence coding on expression for contortrostatin.
- 14. A recombinant DNA molecule according to claim 13, which codes on expression for a contortrostatin protein having at least one biological activity selected from the group consisting of metal binding, proteinase, or disintegrin.

5

PCT/US99/22608

43.

- 15. A purified and isolated DNA molecule consisting essentially of a nucleotide sequence encoding contortrostatin, wherein said nucleotide sequence is selected from the group consisting of:
 - a) nucleotide numbers 1341 to 1535 of SEQ ID NO: 1;
 - b) nucleotide numbers 657 to 1316 of SEQ ID NO: 1;
 - c) nucleotide numbers 87 to 656 of SEQ ID NO: 1;
 - d) nucleotide numbers 87 to 1535 of SEQ ID NO: 1; and
 - e) SEQ ID NO:1.
- 16. A vector comprising a recombinant DNA molecule according to claim 13.
- 17. A host cell transformed with the vector of claim 16, said host cell being selected from the group consisting of animal cells, plant cells, insect cells, yeast and other fungi and bacteria.
- 18. A process for producing contortrostatin comprising the step of culturing the host cell of claim 17.
- 19. The process of claim 18, further comprising the step of recovering the contortrostatin expressed by the host cell.
- 20. A recombinant protein produced by a process comprising the steps of:
 - a) culturing a host cell, said host cell being selected from the group consisting of mammalian cells, plant cells, insect cells, yeast and other fungi and bacteria, said host cell being transformed with a vector, said vector comprising a DNA sequence coding on expression for contortrostatin; and
 - b) recovering the contortrostatin expressed by the host cell.

44

- 21. A pharmaceutically acceptable composition comprising a pharmaceutically acceptable carrier and a purified protein according to claim 1.
- 22. A method of treating a patient having a disease associated with an integrin binding to an integrin receptor, comprising the steps of administering the composition of claim 21 to the patient such that integrin binding to integrin receptors is substantially inhibited and the patient is treated.
- 23. The method of claim 22, wherein the method inhibits platelet aggregation, tumor metastasis, angiogenesis, neovascularization, cell adhesion, invasiveness, or growth.
- 24. The recombinant protein of claim 20, wherein said DNA sequence is selected from the group consisting of:
 - a) nucleotide numbers 1341 to 1535 of SEQ ID NO:1;
 - b) nucleotide numbers 657 to 1316 of SEQ ID NO:1;
 - c) nucleotide numbers 87 to 656 of SEQ ID NO: 1;
 - d) nucleotide numbers 87 to 1535 of SEQ ID NO: 1; and
 - e) SEQ ID NO:1.

1/17

A. Comparison of disintegrin amino acid sequences:

EREDOTOGSPAN<u>FOODAATOKO</u>LFGAQOGEGI.OODQCSFMKKGTI.ORRA**RED**ULDVONGISAGOFRADI.HA ERGEBOOOGSPEN<u>ROODAAIIOI</u>IRPGAQOADGLOODQORFKKKRITIORRA**RGINFIOR**OIGQSADOPRNELYS ERGEDCICKSPAN<u>PCIDARICKL</u>IPGAQCEBELCCIQCSFIEEGIVCRIARGILIDYCNGRSAGCPRNPFH ERGEBOTOCSSPEN<u>ROCDAMICKL</u>RPGAQCABFI.CCDQCKFMKEBSIVC-RARGIIMUYCNGISAGCPRNPFH GRECOCSSPEN<u>ECTRANCE</u>RPGAÇCEBELCCBCCKFSRAGKUCRIP**RAZMPIO**RCIGQSADCPRYH DAPAN ROCIDAATOKI, TTGSQCADGLOCDQCKFMKEGTVC-RARGIDI,-DY-NGISAG---Contortrostatin Albolabrin Trigramin Elegantin Applagin Kistrin

B. Design of PCR primers:

-- Dapan<u>ecedaatckl</u>itgsqcadglccdqckfhkbgtvchrangddl-dy-ngisag-----PCR-1 PCR-2 Agt 10 FORWORD

Agt10 REVERSE

C. Overlapping extension of PCR fragments:

(a) self-extendible molecule:

CM-W (approximately 1300 bp)

S. ---3 ---3 ---3

(b) non-self-extendible molecule:

CH-H (approximately 1300 bp)

3'

-----Agti0 forward

Agt10 reverse

FIG.

PCT/US99/22608

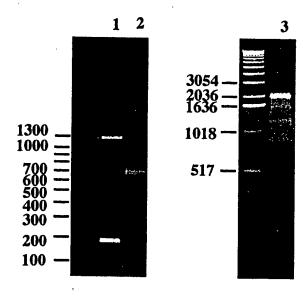


FIG. 2

3/17

5' (BA A	TT C		11 GT C	AA T		_		GC T		GT T		38 TT G		GC A		AG A	GA 7	56 PTG
	CCT	GTC	65 TTC		CCA		CCA				ATG		CAG		101 CTC		GTC	110 AC	
[1]											M	ī	Q	v	L	L	-	т	- [8]
	CTA	TGC	119 TTA	GCA	GCT	128 TTT		TAT			AGC			ATC	155 CTG		тст	164 GGC	
[9]	 L	C	L	 А	 A	F	-	Y	Q	 G	s	s	ī	I	L	 E	s	G	[26]
	AAT	GTT	173 AAT	GAT	TAT		GTA		•		CAA			ACT	209 GCA		ccc	218 AAA	
[27]	N	v	N	D	Y	E	v	L L	Y	P	Q	К	v	T	 А	L	P	ĸ	[44]
	GGA	GCA	227 GTT	CAG	CCA	236 AAG		GAA	245 GAC		ATG		TAT		263 TTT	AAA	GTG	272 AAT	
[45]	G	A	v	Q	P	K	Y	E	D	Т	М	Q	Y	E	F	ĸ	v	N	[62]
	GGA	GAG	281 CCA	GTG	GTC	290 CTT		CTG	299 GAA	AAA	AAT	308 AAA	GGA	CTT	317 TTT	TCA	AAA	326 GAT	
[63]	G	E	P	v	v	L	Н	L	E	ĸ	N	ĸ	G	L	F	s	K	D	- [80]
	TAC	AGC	335 GAG	ACT	CAT		TCC		353 GAT		AGA					AAC		380 CCG	
[81]	Y	s	Е	T	Н	Y	s	s	D	G	R	ĸ	I	т	T	N	P	P	[98]
	GTT	GAG	389 GAT	CAC	TGC	398 TAT	TAT	CAT	407 GGA	CGC	ATC	416 CAG	AAT	GAT	425 GCT	GAC		434 ACT	
[99]	V	E	D	н	С	Y	Y	н	G	R	I	Q	N	D	A	D	s	T	[116
	GCA	AGC	443 ATC	AGT	GCA	452 TGC										CAA			
[117]A	s	I	S	A	С	N	G	L	K	G	Н	F	K	L	Q	G	Е	[134
	ACG	TAC	497 CTT	ATT	GAA	506 CCC			515 CTT	TCC		524 AGT		GCC		GCA			
[135) T	Y	L	I	E	P	L	ĸ	L	s	D	S	E	A	Н	A	V	Y	[152
	AAA 		551 GAA	AAC	GTA	GAA		GAA		GAG		CCC	AAA	ATG	587 TGT				
[153] K				v														[170

4/17

			<u> </u>			1 4		-	~~										
	CAG	_	05 AAT	TGG	GAA	14 TCA	GAT		23 CCC	ATC		32 AAG	GCC	_	41 CAG	TTA	-	50 CTI	
[17]	1)Q	T	N	W	E	s	D	E	P	I	ĸ	ĸ	Α	s	Q	L	N	L	[188]
	ACT	ССТ	659 GAA	CAA	CAA	668 GGA	TTC	CCC	677 CAA	AGA	TAC	686 ATT	GAG	CTT	695 GTT	GTA	GTT	704 GCA	
[189	 9]T	P	E	Q	Q	G	F	P	Q	R	Y	I	E	L	v	v	v	A	[206]
	GAT	CAC	713 AGA	ATG	TTC	722 ACG	AAA	TAC	731 AAC	GGC	AAT	740 TTA	AAT	ACT	749 ATT	AGA	ATA	758 TGG	
[207	 α[τ	н	R	М	F	т	K	Y	N	G	N	L	N	т	I	R	I	W	[224]
	GTA	CAT	767 GAA	CTT	GTC	776 AAC	ACT	ATG	785 AAT	GTG	بلجليل	794 TAC	AGA	ССТ	803 TTG	ልልጥ	יייים	812 CGT	
[225	 5] V		 E	L	 V	N	T			 V	 F	 Y	 R	 P	 L	 N	 I		[242]
			821			830			839			848			857		_	866	• •
	GTC	TCA	CTG	ACT	GAC	CTA	GAA	GTT	TGG	TCA	GAC	CAA	GAT	TTG	ATC	AAC	GTG	CAG	+
[243	3] V	S	L	T	D	L	E	V	W	S	D	Q	D	L	I	N	V	Q	[260]
	CCA	GCA	875 GCG	GCT	GAT	884 ACT	TTG	GAA	893 GCA	TTT	GGA	902 GAC	TGG	AGA	911 GAG	ACA	GTC	920 TTG	
[261	.] P	A	A	A	D	T	L	E	A	F	G	D	W	R	E	T	v	L	[278]
	CTG	AAT	929 CGC	АТА	AGT	938 CAT	GAT	AAT	947 GCT	CAG	TTA	956 CTC	ACG	GCC	965 ATT	GAG	CTT	974 GAT	
[279) L	N	R	I	s	Н	D	N	A	Q.	L	L	т	A	I	E	L	D	[296]
	CCX	C	983	אמשא	aan	992	c cm		.001	000		010	5 00		019			1028	
	 '] G		 T	 I	GGA G	 L	 A	AAC N	AGG R		-			-					
[231	10		1037	1		1046			.055	G	T.	М 1064	С	D	P .073	K	L	5 1082	[314]
	ACA			GTT	CAG					ATA	AAT	CTT		GTT		GTT	ACA		
(315	[] T	G	I	V	Q	D	н	s	A	I					A	,		М	[332]
	GCC		091 GAG	ATG	GGT	100 CAT		CTG							.127 AAT			1136 CAT	
[333] A	Н	E	M	G	н	N	L	G	I	s	Н	D	G	N	Q	С	Н	[350]
	ጥር፥ር		145 GCT	אאר	TCA	154 TGC		_					A G A			ىلىشل		1190 יייייי	
											-								[368]
LODI	., -	ט	-	7.4	ی	_	_	1-1	ټ	E,	E	بد	K	-	V	'n	3	Ľ.	[1000]

5/17

1199 1208 1217 1226 1235 1244 GAG TIT AGC GAT TAT AGT CAG AAT CAG ACT AGG ACT ATT CTT ACT GAT CAT AAC 369]E F S D C S Q N Q Y Q T Y L T D H N [386] 1253 1262 1271 1280 1289 CCA CAA TGC ATG CTC AAT GAA CCC TTG AGA ACA GAT ATT GTT TCA ACT CCA GTT 387]P Q C M L N E P L R T D I V S T P V [404] 1307 1316 1325 1334 1343 1352 TCT GGA AAT GAA CTT TTG GAG ACG GGA GAA AGT GAT GTT GAC GCT CCT GCA 405]S G N B L L E T G E E S D F D A P A [422] 1361 1370 1379 1388 1397 1406 AAT CCG TGC TGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TTT GCA 405]S G N B L L E T G E E S D F D A P A [422] 1361 1370 1379 1388 1397 1406 AAT CCG TGC TGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TTT GCA 423]N P C C D A A T C K L T T G S Q C A [440] GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GAA GAA GAA GTA TCC CGG 441]D G L C C D Q C K F M K E G T V C R [458] 1469 1478 1487 1496 1505 1514 AGA GCA AGG GGT GAT GAC CTG GAT GAT GAT CAA ACA GGC ATA TCT GCT GGC TGT 459]R A R G D D L D D Y C N G I S A G C [476] ACA GGC AGG GTG TTG TGT AAA CAA CAA CAA ATG GAA TGT CAC GCC 477]P R N P F H A * [483] 1571 1586 1592 1564 1550 1559 1568 ACA GGC AGA AAT CCC TTC CAT GCC TAA CAA ACA ATG GAA ATG GAA TGT TCT CCC 477]P R N P F H A * [483] 1691 1691 1690 1649 1658 1667 1676 AGA GAC CCA TGT GTG TTG ATC TGA ATA CAA CAC ATA ATA CAA CCT TGG GCT TCT CTC 1631 1640 1649 1658 1667 1676 AGA GAC CCA TGT GTG TTG ATC TGA ATA CAA CCC TAAT AAT CAA CCT TAG CTT CCT CAA ATC CAA 177]P R N P T H A * [483] 1577 1586 1795 1706 1707 1707 1708 1712 1721 1730 ACA GGC AGT GTG TTG ATC TGA ATA CAA CCT AAT AAT CAA CCT TAG CTT CCT CAA ATC CAA 177]P T T TTC TTC TTC TCC ATA TTT AAT CTA TTT ACC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 ACA GGC ACA TTT TTC TTC TCA CAA ATG CAT TTT AAT CTA TTT ACC TTT TCC TTC TCC 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC CAG CAT TTT AAT 1790 1901 1910 1919 1926 1937 1946 GCA ACA AGT TCT TCC TTT TGA GCT TTT TGC CGT TTC CCC TTC CCC TCC CCC										1015			1006							·
1263																				
1253 1262 1271 1280 1289 1298 1298 CCC CCC CAA TGC CTC AAT GAA CCC TTG AAA ACA GAT ATT GTT TCA ACT CCA GTT 387]P Q C M L N E P L R T D I V S T P V [404] 1307 1316 1325 1334 1343 1352		GAG	TTC	AGC	GAT	TGT	AGT	CAG	AAT	CAA	TAT	CAG	ACA	TAT	CTT	ACT	GAT	CAT	AAC	1
1253 1262 1271 1280 1289 1298 1298 CCC CCC CAA TGC CTC AAT GAA CCC TTG AAA ACA GAT ATT GTT TCA ACT CCA GTT 387]P Q C M L N E P L R T D I V S T P V [404] 1307 1316 1325 1334 1343 1352																				
CCA CAA TGC ATG CTC AAT GAA CCC TTG AGA ACA GAT ATT GTT TCA ACT CCA GTT 387]P Q C M L N E P L R T D I V S T P V [404] 1307	[369) E	F	S	D	C	S	Q	N	Q	Y	Q	T	Y	L	T	D	H	N	[386]
CCA CAA TGC ATG CTC AAT GAA CCC TTG AGA ACA GAT ATT GTT TCA ACT CCA GTT 387]P Q C M L N E P L R T D I V S T P V [404] 1307																				
CCA CAA TGC ATG CTC AAT GAA CCC TTG AGA ACA GAT ATT GTT TCA ACT CCA GTT 387]P Q C M L N E P L R T D I V S T P V [404] 1307				1253		:	1262			1271			1280			1289			1298	
387]P Q C M L N E P L R T D I V S T P V [404] 1307		CCA	CAA	TGC	ATG													CCA	GTT	
TOT GGA AAT GAA CTT TTG GAG ACG GGA GAA GAA AGT GAC TTT GAC GCT CCT GCA 405]S G N E L L E T G E E S D F D A P A [422] AAT CCG TGC TGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TGT GCA 423]N P C C D A A T C K L T T G S Q C A [440] GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GGG TCA CAG TGT GCA 423]N P C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAA GGA ACA GGG TCA CAG TGC TGC 441]D G L C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAA TTC GCT GCC TGC 459]R A R G D D L D D Y C N G I S A G C [476] 1523																				
TOT GGA AAT GAA CTT TTG GAG ACG GGA GAA GAA AGT GAC TTT GAC GCT CCT GCA 405]S G N E L L E T G E E S D F D A P A [422] AAT CCG TGC TGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TGT GCA 423]N P C C D A A T C K L T T G S Q C A [440] GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GGG TCA CAG TGT GCA 423]N P C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAA GGA ACA GGG TCA CAG TGC TGC 441]D G L C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAA TTC GCT GCC TGC 459]R A R G D D L D D Y C N G I S A G C [476] 1523	[385	ם וי	0	C	м	т.	NT	10	D	т.	Ð	T	ъ	-	3.7		TT.	п	17	[404]
TCT GGA AAT GAA CTT TTG GAG ACG GGA GAA AGT GAA AGT GAC TTT GAC GCT CCT GCA 405]S G N E L L E T G E E S D P D A P A [422] 1361 1370 1379 1388 1397 1406 AAT CCG TGC TGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TGT GCA 423]N P C C D A A T C K L T T G S Q C A [440] 1415 1424 1433 1442 1451 1460 GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GAA GAA ACA GTA TGC CGG 441]D G L C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT GAC ACA ACA GGG TCA CAG TGT GCC TGC 459]R A R G D D L D D Y C N G I S A G C [476] CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG TAG TCT GCA GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TGT ATC TGA ATC CAG ACA ATA ATC CAC CTT GCT GCT TCT ACA GGC AGT GTG TGT ATC TGA ATC CAG ACA ATA ATC CAC CTT CTT CTC 1631 1640 1649 1658 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA CTA GTA AAT CAC CTT TAG CTT CCA GAT GGT AGA GAC CCA TCT GCC TGC ATC CTA CTA CTA GTA AAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA AAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TAT TAAT CAC CTT TAG CTT TCC TGT 1739 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTT ACT GCT 1847 1856 1865 1865 1874 CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946	(50)	12	¥	C	1.1		14	-	E		. **	•	D	_	٧	3	-	F	٧	[404]
TCT GGA AAT GAA CTT TTG GAG ACG GGA GAA AGT GAA AGT GAC TTT GAC GCT CCT GCA 405]S G N E L L E T G E E S D P D A P A [422] 1361 1370 1379 1388 1397 1406 AAT CCG TGC TGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TGT GCA 423]N P C C D A A T C K L T T G S Q C A [440] 1415 1424 1433 1442 1451 1460 GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GAA GAA ACA GTA TGC CGG 441]D G L C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT GAC ACA ACA GGG TCA CAG TGT GCC TGC 459]R A R G D D L D D Y C N G I S A G C [476] CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG TAG TCT GCA GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TGT ATC TGA ATC CAG ACA ATA ATC CAC CTT GCT GCT TCT ACA GGC AGT GTG TGT ATC TGA ATC CAG ACA ATA ATC CAC CTT CTT CTC 1631 1640 1649 1658 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA CTA GTA AAT CAC CTT TAG CTT CCA GAT GGT AGA GAC CCA TCT GCC TGC ATC CTA CTA CTA GTA AAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA AAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TAT CAC CTT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TAT TAAT CAC CTT TAG CTT TCC TGT 1739 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTT ACT GCT 1847 1856 1865 1865 1874 CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946				1207			1216			1225			1 2 2 4							
1361		_~~	-																	
1361 1370 1379 1388 1397 1406 AAT CCG TGC GGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TGT GCA 423]N P C C D A A T C K L T T G S Q C A [440] 1415 1424 1433 1442 1451 1460 GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GAA GGA ACA GTA TGC CGG 441]D G L C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAT GGC ATA TCT GCT GCT TGT 459]R A R G D D L D D Y C N G I S A G C [476] 1523 1532 1541 1550 1559 1568 CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TGT ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTC CAC TCC CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCC CAA ATC CAA 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TC CTC CTA CTA GTA AAT CAC CTT TAG CTT CCC GTT TGT CTC 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TCC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA AAA AAA AAA AAA AAA AA		TCT	GGA	AAT	GAA	CTT	TIG	GAG	ACG	GGA	GAA	GAA	AGT	GAC	.III.	GAC	GCT	CCT	GCA	
1361 1370 1379 1388 1397 1406 AAT CCG TGC GGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TGT GCA 423]N P C C D A A T C K L T T G S Q C A [440] 1415 1424 1433 1442 1451 1460 GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GAA GGA ACA GTA TGC CGG 441]D G L C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAT GGC ATA TCT GCT GCT TGT 459]R A R G D D L D D Y C N G I S A G C [476] 1523 1532 1541 1550 1559 1568 CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TGT ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTC CAC TCC CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCC CAA ATC CAA 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TC CTC CTA CTA GTA AAT CAC CTT TAG CTT CCC GTT TGT CTC 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TCC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA AAA AAA AAA AAA AAA AA	_																			
AAT CCG TGC TGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TGT GCA 423]N P C C D A A T C K L T T G S Q C A [440] 1415	[405]S	G	N	E	L	L	E	T	G	E	E	S	D	F	D	Α	P	Α	[422]
AAT CCG TGC TGC GAT GCT GCA ACA TGT AAA CTG ACA ACA GGG TCA CAG TGT GCA 423]N P C C D A A T C K L T T G S Q C A [440] 1415																				
423]N P C C D A A T C K L T T G S Q C A [440] 1415			1	1361			1370			1379		:	1388			1397		:	1406	
423]N P C C D A A T C K L T T G S Q C A [440] 1415		AAT	CCG	TGC	TGC	GAT	GCT	GCA	ACA	TGT	AAA	CTG	ACA	ACA	GGG	TCA	CAG	TGT	GCA	
1415 1424 1433 1442 1451 1460 GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GAA GGA ACA GTA TGC CGG 441]D G L C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAT GGC ATA TCT GCT GGC TGT 459]R A R G D D L D D Y C N G I S A G C [476] 1523 1532 1541 1550 1559 1568 CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1786 ATC CAA ATT CTG TCA TTC TCC ATA TTT AAT CTA TTT ACC TTT TCC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA ACC CTC ATG GCA TTA CAA ACC TTT TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA AAA GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946																				
1415 1424 1433 1442 1451 1460 GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GAA GGA ACA GTA TGC CGG 441]D G L C C D Q C K F M K E G T V C R [458] AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAT GGC ATA TCT GCT GGC TGT 459]R A R G D D L D D Y C N G I S A G C [476] 1523 1532 1541 1550 1559 1568 CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1786 ATC CAA ATT CTG TCA TTC TCC ATA TTT AAT CTA TTT ACC TTT TCC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA ACC CTC ATG GCA TTA CAA ACC TTT TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA AAA GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946	[423] N	P	Ċ	С	D	A	A	т	С	K	L	т	T	G	s	0	C	A	[440]
GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GAA GGA ACA GTA TGC CGG 441]D G L C C D Q C K F M K E G T V C R [458] 1469	•	•				. –						_	-	•	_	-	*	_	••	(,
GAT GGA CTG TGT TGT GAC CAG TGC AAA TTT ATG AAA GAA GGA ACA GTA TGC CGG 441]D G L C C D Q C K F M K E G T V C R [458] 1469			-	1415		-	1424			1433			1442			1/51			1460	
441]D G L C C D Q C K F M K E G T V C R [458] 1469		ርልጥ	-		ጥርጥ															
1469 1478 1487 1496 1505 1514 AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAT GGC ATA TCT GCT GGC TGT 459]R A R G D D L D D Y C N G I S A G C [476] CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT		OV I			191	191			100		T T T	A1G	AAA	UAA	GGA	ACA	GIA	IGC	نافات	
1469 1478 1487 1496 1505 1514 AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAT GGC ATA TCT GCT GGC TGT 459]R A R G D D L D D Y C N G I S A G C [476] CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT	[4 4 9	15	~~~						~~											3
AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAT GGC ATA TCT GCT GGC TGT 459]R A R G D D L D D Y C N G I S A G C [476] 1523	[441	ם נ	G	ъ	Ç	C	ע	Q	C	K	F.	M	K	E	G	Т	V	С	R	[458]
AGA GCA AGG GGT GAT GAC CTG GAT GAT TAC TGC AAT GGC ATA TCT GCT GGC TGT 459]R A R G D D L D D Y C N G I S A G C [476] 1523																				
459]R A R G D D L D D Y C N G I S A G C [476] CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA GCA 477]P R N P F H A * [483] 1577			_			_														
1523 1532 1541 1550 1559 1568 CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946		AGA	GCA	AGG	GGT	GAT	GAC	CTG	GAT	GAT	TAC	TGC	AAT	GGC	ATA	TCT	GCT	GGC	TGT	
1523 1532 1541 1550 1559 1568 CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA GCA 477]P R N P F H A * [483] 1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946																				
CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA GCA 477] P R N P F H A * [483] 1577	[459] R	Α	R	G.	D	D	L	D	D	Y	C	N	G	I	S	Α	G	C	[476]
CCC AGA AAT CCC TTC CAT GCC TAA CCA ACA ATG GAG ATG GAA TGG TCT GCA GCA 477] P R N P F H A * [483] 1577																				Ì
1577			1	1523		1	1532		:	1541		1	L550		-	1559			1568	
1577		CCC	AGA	AAT	CCC	TTC	CAT	GCC	TAA	CCA	ACA	ATG	GAG	ATG	GAA	TGG	TCT	GCA	GCA	į
1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT																				
1577 1586 1595 1604 1613 1622 ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631 1640 1649 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT	[477	1P	R	N	P	' F	н	Α	*	[483]	1									
ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631	• • • •	•	•-	•	_	_					,									
ACA GGC AGT GTG TTG ATC TGA ATA CAG CCT AAT AAT CAA CCT CTG GCT TCT CTC 1631			1	577		-	596			1505		-	604		-	612		-	1622	
1631 1640 1649 1658 1667 1676 AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT		א מיא			ama							_			_					i
AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685		лсл	GGC	MOI	G10	110	AIC	IGA	VIV	CAG	CCI	AAI	AAI	CAA	CCI	CIG	GCT	ICI	CIC	
AGA TTT GAT CAT GGA GAT CCT TCT TCC AGA AGG TTT CAC TTC CCT CAA ATC CAA 1685				<i>-</i> 233		_	- 40					_			_			_		
1685 1694 1703 1712 1721 1730 AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT			_									_								
AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739		AGA	TTT	GAT	CAT	GGA	GAT	CCT	TCT	TCC	AGA	AGG	TTT	CAC	TTC	CCT	CAA	ATC	CAA	
AGA GAC CCA TCT GCC TGC ATC CTA CTA GTA AAT CAC CCT TAG CTT CCA GAT GGT 1739																				
1739 1748 1757 1766 1775 1784 ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946																				ĺ
ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793		AGA	GAC	CCA	TCT	GCC	TGC	ATC	CTA	CTA	GTA	AAT	CAC	CCT	TAG	CTT	CCA	GAT	GGT	
ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793																				
ATC CAA ATT CTG TAA TAT TTC TTC TCC ATA TTT AAT CTA TTT ACC TTT TGC TGT 1793			1	739		1	748		1	1757		1	1766		3	1775		1	1784	- 1
1793 1802 1811 1820 1829 1838 AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946		ATC																		1
AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847																				
AAC AAA ACC TTT TTC CTG TCA CAA AGC TCC ATG GGC ATG TAC AGC TTA TCT GCT 1847			1	793		1	802		1	1811	•	1	820		1	829		1	1838	
1847 1856 1865 1874 1883 1892 GTC AAG AAA AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946		AAC																		j
GTC AAG AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946																				l
GTC AAG AAA AAT GGC CAT TTT ACC GTT TGC CAG TTA CAA AGC ACA TTT AAT 1901 1910 1919 1928 1937 1946			1	847		1	856		-	865		4	874		-	202		-	202	
1901 1910 1919 1928 1937 1946		כייירי																		
		310	AAG	AAA	AAA	AAT	GGC	CHT	I I I	ACC	GIT.	IGC	CAG	TTA	CAA	AGC	ACA	1.1.1,	AA'I'	
			_			_			_				 -							
GCA ACA AGT TCT TCC TTT TGA GCT GAT GTA TTC AAA GTC AAT GCT TCC TCT CCC																				
		GCA	ACA	AGT	TCT	TCC	TTT .	TGA	GCT	GAT	GTA	TTC	AAA	GTC	AAT	GCT	TCC	TCT	CCC	
												•								

PCT/US99/22608

6/17

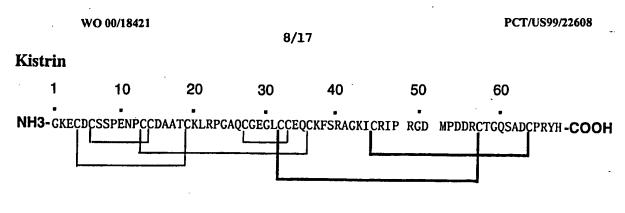
1955 1964 1973 1982 1991 2000 AAA ATT TCA TGC TGG CTT CCC AAG ATG TAG CTG CTT CCG TC<u>A ATA AA</u>C AAA CTA

Proprotein domain:

7/17

LTAIELDGETIGLANRGTMCDPRLSTGIVQDHSAINLMVAVTMAHEMGHRIGISHDGNQCHCDANSCINSEELAEQLSFEFSDCSQNQYQTYLTDHNPQCMLAEPLAFDIVSTPVSGNELL. LTATIFNGNVIGRAPVGGMCDPRRSVAIVRDHNAIVFVVAVTMT<u>HEMGHNIGMH</u>HDEDKCNCN. TCINSKVLSRQPSKYFSECSKDYYQTFLTNHNPQCILANPLAFTDTVSTPVSGNELL. LTAIDL. DRVIGLAYVGSHCHPRRSTGIIQDYSEINLVVAVIMA<u>HEMGHRIGINH</u>DSGYCSCGDYACIRRPEISPEPSFFFSNCSYFECMDFINNPPECILAMEPLGTDIISPPVCGNELL. ltaldengptigyayigsmchprrygolyqdyspinlvvavima<u>hemghnlgihh</u>dtgscscgdypcihgptisnepskffsncsyi**qcmd**fimhnpeciineplctdisppvcgnel LTSIAFDEQIIGRAYIGGICDPKRSTGVVQDHSEINLAVAVTMTHELGHNLAIHHDTDSCSCGGYSCIHSPVISDEPSKYFSDCSYIQCMEFIHNQKPQCILKKPLRTDTVSTPVSGNELI eqoge . Pory ielavvadhrmftryngallatiri whelavtprplairi rysltdewsdodlinvopaaadtleafgd . Mettvllari shdaaol eqore . Poryikijifudhgwytkysgnseritkruhqminninnmcralmivytlsvleimsekdlituq. Asapytlilegamretyllaryshdhaol ehokynpfrevelælvudkamvtkamgdldri ktrmyeivntvmeiyrymy i hvalvgleimsnedki tvkpeagytlaa . Fgemrttdllprkkhdnaol eqorydfyryi effuvudggtuttanngdldrikarhyelanivnei frylymhyalvgleimsngdriturpdudytlas . Faemrktdllyrkhdnaldl eho.....ryvelfivvdhgmytkyngdsdkirqrvhqmvnimkesytymyidilllagieimsngdlinvqpaspntlns.fgemretdllrrkshdaadl hiqvilyticiaappygssiiilesgnyndyevlypokvtalpkgavqpryedthqyepryngepvylhiernkglfskdysethyssdgriitynppye miqulliticlavf pyggssiilesgalladyevvpekvtalpkgavqqkyedaaqyefkvagepvulhlernkglf sedyseihyspdreitaypsve MIGVILVTI CLAAF PYQGSSI I LÆSGAVNDYEVI Y PRKVTAL PKGAVQPKYEDAMGYELKVNGEPVVLHIGNOKGLFSKDYSETHY SPDGREI TTYPLNE atrpkgavopkyedamoyeprvngepvalalennkglpskdyseihyspdcreiftyppvr miqvilivticlaafpyqgssiilesgnvndyeviyprkvtalpkgavqpryedthqyelkvngepvvlhlernkglfskdysethysfdgrittynpsvb NGQPCLDNYGYCYNGNCP INYHQCYDLFGADVYEAEDSCFERNQKGNYYGYCRKENGNKI PCAPEDVKGGRLYCKDNSPGQNNPCKMF YSNEDEHKGNYL NGOPCLDYYGYCYNGNCPIMYHQCYALFGADVYEAEDSCFKDNQKGNYYGYCRKENGKKIPCAPEDVKCGRLYCKDNSPGGRNPCKMFYSNDDEHKGAYL dhcyyhgri qwdadstasi sachglkghfkilggetyli epilklsdseahavykyenvekedeapkhcgvtqtnwesdepikka sqlalitp dhcyyhgri endadstasisacnglkghfklaggenyll eplkl.Pdseahavykyenvekedealkhcgvtg. nwesyepikkasglvvta dhcyyhgriendadstasisachgyfkuqrettyieplalhadseahayfkyenvekederprocytq.nmksyepikasqlafta dhcyyhgri endadsta si sacdglkghfklagemyl i eplelsdseahavfky envekedeppkacgvtq . Nwesy estkrasqlavyp DHCYYHGRIENDADSTASISACNGLKGHFKLQGEMYLIEPLKLSDSEAHAVFKLKNVEKEDEAPNACGYTQ. NMESYEPIKKASDLAIANP etgeesdf----dapanpccdaatcklitgsqcadglccdqckfwkegtvcrra<u>rgd</u>d. Dlddycngi sagcprnpfha• eagiecdgsle...npccyattckarpgsqcaeglccdqcrfmkkgtvcrvsmydrn.ddtctqqsadcprnglyg EAGEDCDCGSPA...NPCCDAATCKLIPGAQCGEGLCCDQCSFIEEGTVCRIARGD.DLDDYCNGRSAGCPRNPFHA evgeecdcgtpencqneccdaatcklasgsqcghgdcceqckfsksgtecrashsecdpaehctgqssecpadvfhk evgeecdcgtpencqneccdaatcklasgsqcghgdcceqcafsksgtecrasmsecdpaehctgqssecPadvfhr 360 550 350 340 240 330 140 440 320 PGTKCADGKVCSNGHCVDVATAY. **PGTKCADGKVCSNGHCVDVATAY** 310 200 900 200 490 Metalloproteinase domain: Disintegrin domain: C-terminal domain: Jararhagin Jararhagin Jararhagin Jararhagin Jararhagin Jararhagin Jararhagin Trigramin frigremin Trigramin Trigramin Trigramin Ht-e g Cat Gat Sat Cat 3 3 3

FIGURE 4



Contortrostatin

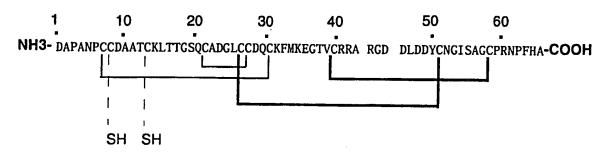
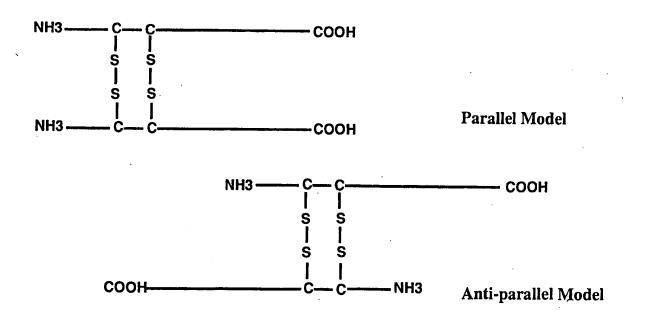
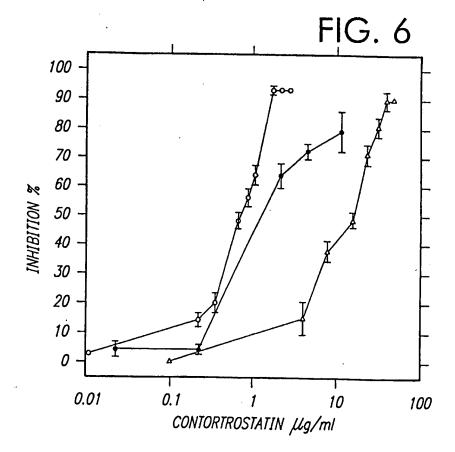


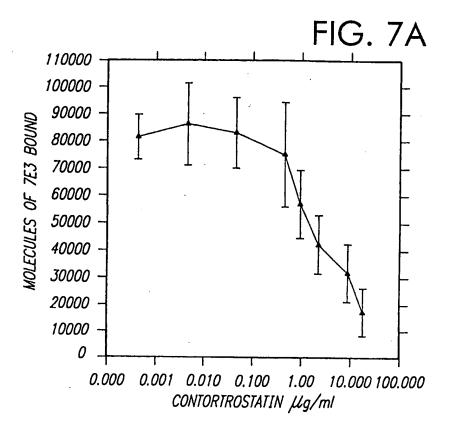
FIGURE 5B

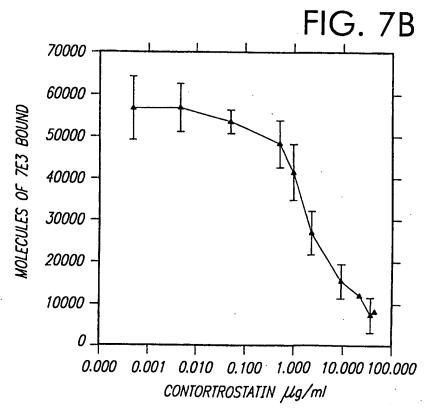


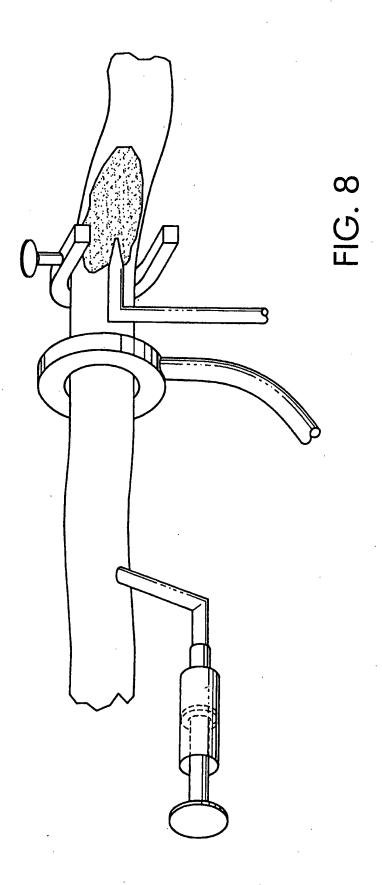
PCT/US99/22608



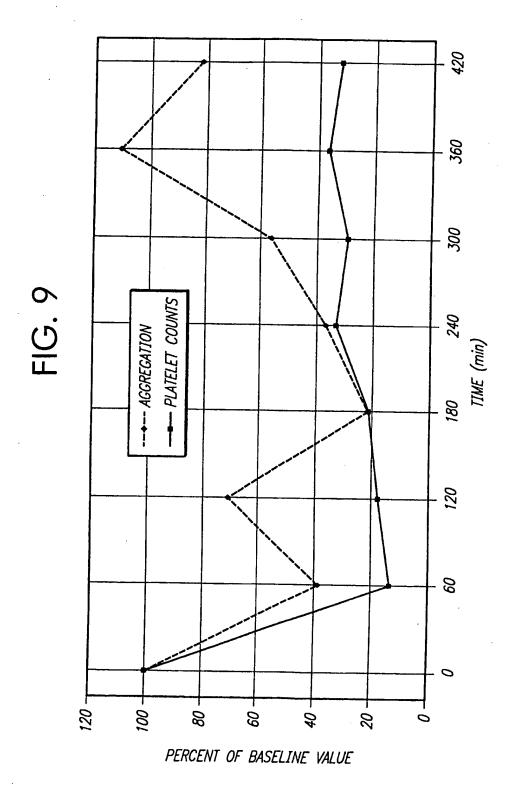
PCT/US99/22608



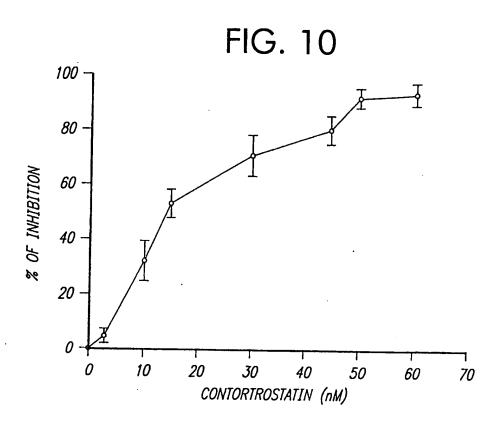


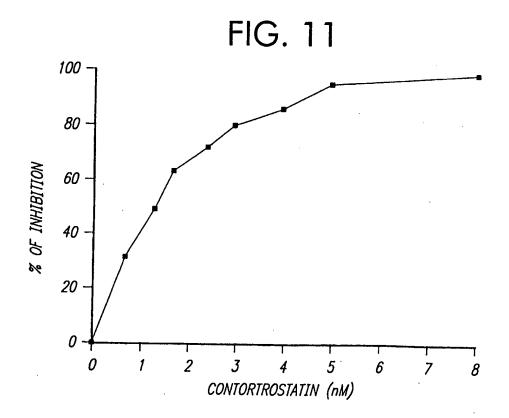


PCT/US99/22608

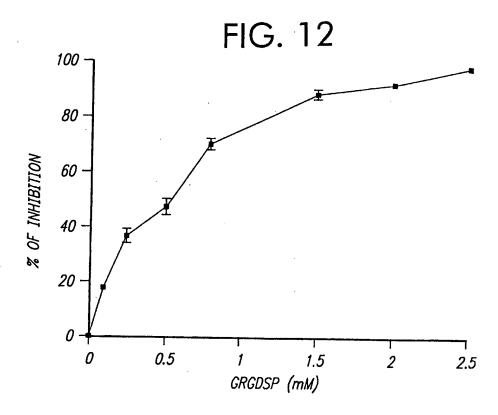


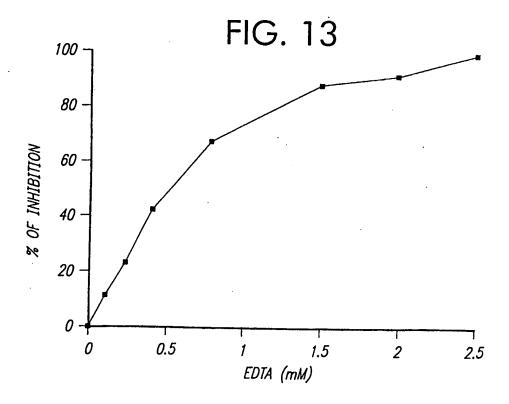
PCT/US99/22608





PCT/US99/22608





15/17

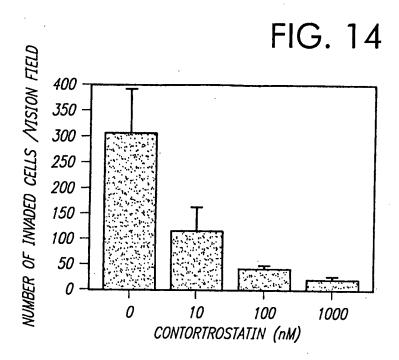


FIG. 15

FIG. 15

FIG. 15

FIG. 15

FIG. 15

PCT/US99/22608



FIG. 16A

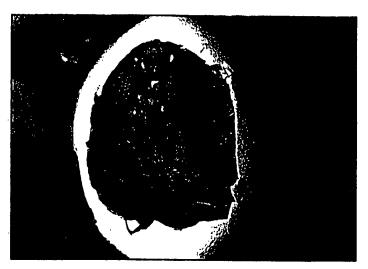


FIG. 16B

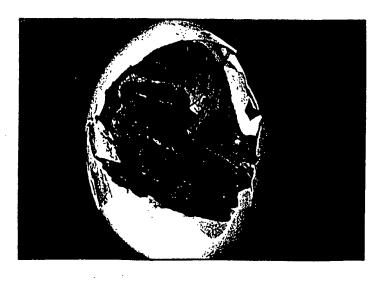
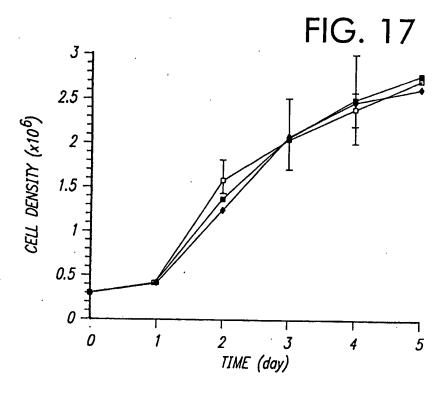
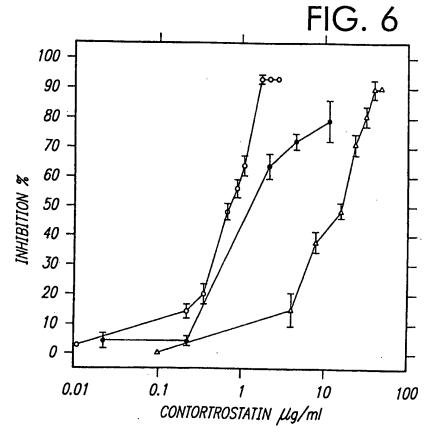


FIG. 16C

PCT/US99/22608





1

SEQUENCE LISTING

<110> Markland, Francis S. Zhou, Qing

<120> Contortrostatin (CN) and Methods for its Use In Preventing Metastasis and Other Conditions

<130> 1920-338C3

<140> N/A

<141> 1999-09-29

<150> US 09/163,047

<151> 1998-09-29

<160> 6

<170> PatentIn Ver. 2.0

<210> 1

<211> 2029

<212> DNA

<213> Agkistrodon contortrix

<400> 1

gaattcgggg tcaatagagg aagagctcaa gttggcttga aagcaggaag agattgcctg 60 tottocagoo aaatocagoo gocaaaatga tocaggttot ottggtgact otatgottag 120 cagettttee ttateaaggg agetetataa teetggaate tgggaatgtt aatgattatg 180 aagtactgta tccacaaaaa gtcactgcat tgcccaaagg agcagttcag ccaaagtatg 240 aaqacaccat qcaatatgaa tttaaagtga atqqaqaqcc aqtqqtcctt cacctggaaa 300 aaaataaagg acttttttca aaagattaca gcgagactca ttattcctct gatggcagaa 360 aaattacaac aaaccctccg gttgaggatc actgctatta tcatggacgc atccagaatg 420 atgctgactc aactgcaagc atcagtgcat gcaacggttt gaaaggacat ttcaagcttc 480 aaggggagac gtaccttatt gaacccttga agctttccga cagtgaagcc catgcagtct 540 acaaatatga aaacgtagaa aaagaagatg aggcccccaa aatgtgtggg gtaacccaga 600 ctaattggga atcagatgag cccatcaaaa aggcctctca gttaaatctt actcctgaac 660 aacaaggatt cccccaaaga tacattgagc ttgttgtagt tgcagatcac agaatgttca 720 cgaaatacaa cggcaattta aatactatta gaatatgggt acatgaactt gtcaacacta 780 tgaatgtgtt ttacagacct ttgaatattc gtgtctcact gactgaccta gaagtttggt 840 cagaccaaga tttgatcaac gtgcagccag cagcggctga tactttggaa gcatttggag 900 actggagaga gacagtettg etgaategea taagteatga taatgeteag ttaeteaegg 960 ccattgagct tgatggagaa actataggat tggctaacag gggcaccatg tgcgacccga 1020 agetttetae aggaattgtt caggateata gtgeaataaa tetttgggtt geagttaeaa 1080 tgccccatga gatgggtcat aatctgggta ttagtcacga tggaaatcag tgtcattgcg 1140 atgctaactc atgcattatg agtgaagaac taagagaaca actttccttt gagttcagcg 1200 attgtagtca gaatcaatat cagacatatc ttactgatca taacccacaa tgcatgctca 1260 atgaaccett gagaacagat attgtttcaa etceagttte tggaaatgaa ettttggaga 1320

2

cyggagaaga aagtgacttt gacgctcctg caaatccgtg ctgcgatgct gcaacatgta 1380
aactgacaac agggtcacag tgtgcagatg gactgtgttg tgaccagtgc aaatttatga 1440
aagaaggaac agtatgccgg agagcaaggg gtgatgacct ggatgattac tgcaatggca 1500
tatctgctgg ctgtcccaga aatcccttcc atgcctaacc aacaatggag atggaatggt 1560
ctgcagcaac aggcagtgtg ttgatctgaa tacagcctaa taatcaacct ctggcttctc 1620
tcagatttga tcatggagat ccttcttcca gaaggtttca cttccctcaa atccaaagag 1680
acccatctgc ctgcatccta ctagtaaatc acccttagct tccagatggt atccaaattc 1740
tgtaatattt cttctccata tttaatctat ttaccttttg ctgtaacaaa accttttcc 1800
tgtcacaaag ctccatgggc atgtacagct tatctgctgt caagaaaaaa aatggccatt 1860
ttaccgtttg ccagttacaa agcacattta atgcaacaag ttcttccttt tgagctgatg 1920
tattcaaagt caatgcttcc tctcccaaaa tttcatgctg gcttcccaag atgtagctgc 1980
ttccgtcaat aaacaaacta ttctcattca aaaaaaaaa cccgaattc 2029

<210> 2

<211> 483

<212> PRT

<213> Agkistrodon contortrix

<400> 2

Met Ile Gln Val Leu Leu Val Thr Leu Cys Leu Ala Ala Phe Pro Tyr

1 5 10 15

Gln Gly Ser Ser Ile Ile Leu Glu Ser Gly Asn Val Asn Asp Tyr Glu 20 25 30

Val Leu Tyr Pro Gln Lys Val Thr Ala Leu Pro Lys Gly Ala Val Gln
35 40 45

Pro Lys Tyr Glu Asp Thr Met Gln Tyr Glu Phe Lys Val Asn Gly Glu
50 55 60

Pro Val Val Leu His Leu Glu Lys Asn Lys Gly Leu Phe Ser Lys Asp 65 70 75 80

Tyr Ser Glu Thr His Tyr Ser Ser Asp Gly Arg Lys Ile Thr Thr Asn 85 90 95

Pro Pro Val Glu Asp His Cys Tyr Tyr His Gly Arg Ile Gln Asn Asp 100 105 110

Ala Asp Ser Thr Ala Ser Ile Ser Ala Cys Asn Gly Leu Lys Gly His 115 120 125

Phe Lys Leu Gln Gly Glu Thr Tyr Leu Ile Glu Pro Leu Lys Leu Ser 130 135 140

3

Asp Glu Ala Pro Lys Met Cys Gly Val Thr Gln Thr Asn Trp Glu Ser 170. Asp Glu Pro Ile Lys Lys Ala Ser Gln Leu Asn Leu Thr Pro Glu Gln 185 Gln Gly Phe Pro Gln Arg Tyr Ile Glu Leu Val Val Ala Asp His 195 200 Arg Met Phe Thr Lys Tyr Asn Gly Asn Leu Asn Thr Ile Arg Ile Trp 215 Val His Glu Leu Val Asn Thr Met Asn Val Phe Tyr Arg Pro Leu Asn 230 235 Ile Arg Val Ser Leu Thr Asp Leu Glu Val Trp Ser Asp Gln Asp Leu 245 250 Ile Asn Val Gln Pro Ala Ala Ala Asp Thr Leu Glu Ala Phe Gly Asp 260 265 Trp Arg Glu Thr Val Leu Leu Asn Arg Ile Ser His Asp Asn Ala Gln 275 280 Leu Leu Thr Ala Ile Glu Leu Asp Gly Glu Thr Ile Gly Leu Ala Asn 295 300 Arg Gly Thr Met Cys Asp Pro Lys Leu Ser Thr Gly Ile Val Gln Asp 305 310 315 His Ser Ala Ile Asn Leu Trp Val Ala Val Thr Met Ala His Glu Met 325 . 330 Gly His Asn Leu Gly Ile Ser His Asp Gly Asn Gln Cys His Cys Asp 340 345 Ala Asn Ser Cys Ile Met Ser Glu Glu Leu Arg Glu Gln Leu Ser Phe Glu Phe Ser Asp Cys Ser Gln Asn Gln Tyr Gln Thr Tyr Leu Thr Asp 375 380 His Asn Pro Gln Cys Met Leu Asn Glu Pro Leu Arg Thr Asp Ile Val 385 390 395 Ser Thr Pro Val Ser Gly Asn Glu Leu Leu Glu Thr Gly Glu Glu Ser 405 4:10

4

Asp Phe Asp Ala Pro Ala Asn Pro Cys Cys Asp Ala Ala Thr Cys Lys 420 425 Leu Thr Thr Gly Ser Gln Cys Ala Asp Gly Leu Cys Cys Asp Gln Cys 440 Lys Phe Met Lys Glu Gly Thr Val Cys Arg Arg Ala Arg Gly Asp Asp 450 455 Leu Asp Asp Tyr Cys Asn Gly Ile Ser Ala Gly Cys Pro Arg Asn Pro 470 475 480 Phe His Ala <210> 3 <211> 21 <212> DNA <213> Trimeresurus gramineus <400> 3 gtttacaggt tgcagcatcg c 21 <210> 4 <211> 21 <212> DNA <213> Trimeresurus gramineus <400> 4 gcgatgctgc aacctgtaaa c 21 <210> 5 <211> 21 <212> DNA <213> Bacteriophage Lambda gt10 <400> 5 gcgatgctgc aacctgtaaa c 21 <210> 6 <211> 24 <212> DNA <213> Bacteriophage Lambda gt10 <400> 6 cttatgagta tttcttccag ggta 24

INTERNATIONAL SEARCH REPORT International application No. PCT/U399/22608 CLASSIFICATION OF SUBJECT MATTER IPC(7) :A61K 38/00, 38/16; C07K 14/00 US CL : 514/12; 530/324 According to International Patent Classification (IPC) or to both national classification and IPC **FIELDS SEARCHED** Minimum documentation searched (classification system followed by classification symbols) 514/12; 530/324 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) DOCUMENTS CONSIDERED TO BE RELEVANT Category* Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. Database CAPLUS, Accession No. 1994:572676, CLARK et al. 1, 2, 6, 12, 21 'Structurally distinct disintegrins contorteostatin and multisquamatin Y differentially regulate platelet tyrosine kinase phosphorylation, abstract, J. Biol. Chem. 1994, Vol. 269, pages 21940-21943, see abstract. X Database CAPLUS, Accession No. P90457, CHIRON CORP., 'DNA encoding snake venom fibrolase - for use as thrombolytic agent,' abstract, EP 323722 A, 12 July 1989, see sequence printout. Further documents are listed in the continuation of Box C. See patent family annex. later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance earlier document published on or after the international filing date ۰E document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step document which may throw doubts on priority claim(s) or which is cited to establish the publication thate of another citation or other special reason (as specified) ·L when the document is taken alone document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination ·0· document referring to an oral disclosure, use, exhibition or other being obvious to a person skilled in the art document published prior to the international filing date but later than document member of the same patent family the priority date claimed Date of the actual completion of the international search Date of mailing of the international search report 0 8 MAR 2000 03 FEBRUARY 2000 Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Authorized officer MICHAEL BORIN Washington, D.C. 20231 Facsimite No. (703) 305-3230 (703) 308-0196 Telephone No.

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/22608

	"	CT/US99/2260	76		
C (Continua	ation). DOCUMENTS CONSIDERED TO BE RELEVANT				
Category*	Citation of document, with indication, where appropriate, of the relevant p	assages	Relevant to claim No.		
X	Database CAPLUS, Accession No. R06494, BIOGEN INC platelet activation inhibiting polypeptide from snake venor for preventing agglutination and release in vivo or vitro ar recombinant DNA encoding it, abstract, WO 9008772 A, August 199), see sequence printout.	n - used id new			
x	Database CAPLUS, Accession No. W14083, MOGAM BIOTECHNOLOGY RES. INST., 'Peptide derived from K salmosa viper venom- useful as blood platelet aggregation inhibitor, for the management of thrombosis,' abstract, FR A1, 10 January 1997, see sequence printout.		1, 7		
			·		
			·		
			·		

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/22608

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)
This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:
Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
·
Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)
This International Searching Authority found multiple inventions in this international application, as follows:
Please See Extra Sheet.
As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos
4. X No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-7, 12, 21; species A
Remark on Protest The additional search fees were accompanied by the applicant's protest.
No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet(1))(July 1992)★

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/22608

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional examination fees must be paid.

- Group I, claims 1-7,12,21,drawn to contortrostatin and its fragments, and their compositions.
- II. Claim 8, drawn to polypeptide recognized by an antibody to contortrostatin.

Claims 13-15 drawn to recombinant DNA molecule.

Claims 16, 17, drawn to vector, and a host cell.

Claims 18, 19, drawn to method of making.

III.

IV.

VI. Claim 22, 23, drawn to method of treatment.

VII. Claims 9-11, 20, 24, drawn to recombinant protein produced from DNA of SEQ ID No. 1,or from its fragments.

The inventions listed as Groups I-IV do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack in the or corresponding special technical features for the following reasons: The inventions listed as Groups I. II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Group I is the technical feature that links Groups I to VII. Group I is not the contribution over the prior art because it is suggested or taught by references teaching contortrostatin or its biologically active variants. For example, such references are the following: FR2736266, or WO9008772, or EP323722, or Clark et al. (J. Biol. Chem. (1994), 269(35), 21940-3, 1994). Therefore, the lack of unity is present because the linking technical feature is not a "special technical feature" as defined by PCT Rule 13.2.

The products of Groups I and VII are distinct because although they are related as disclosed, they are capable of separate manufacture, use, or sale as claimed. The products of group I can be made by a different way, e.g., by peptide synthesis, which is not dependent upon use of DNA of SEQ ID 1 or fragments thereof.

The polypeptides of Groups I, VII DNA of Group III, vector and host cell of Group IV are drawn to independent and/or patentably distinct compounds since each of these compounds possess different structure (e.g., primary, secondary and tertiary structure) and/or physico-chemical properties, and/or capable of separate manufacture and/or use.

Inventions I and VI are related as product and process of use. The method of use can be practiced with a broad variety of polypeptides other than contortrostatin, e.g., other RGD-containing peptides.

Inventions V and I are related as process of making and product made. The product of Group I can be obtained using other methods such as peptide synthesis.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack unity of invention because they are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for more than one species to be examined, the appropriate additional examination fees must be paid.

In regard to Group I, the species are the peptides having the following residues of SEQ ID 2:

Group A: residues 419 to 483;

Group B: residues 191-410;

Group C: residues 1-190.

The sequences do not have a common core structure required for common utility.

In regard to Group III, the species are the nucleotides having the following residues of SEQ ID 1:

Group D: residues 1341-1535;

Group E: residues 657-1316;

Group F: residues 87-656

Group F: residues 87-1535

the sequences do not have a common core structure required for common utility.

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:
BLACK BORDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
☐ FADED TEXT OR DRAWING
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
SKEWED/SLANTED IMAGES
COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
Потикр.

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.